(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 September 2003 (25.09.2003)

PCT

(10) International Publication Number
WO 03/077859 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number: PCT/US03/07727

(22) International Filing Date: 12 March 2003 (12.03.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/363,807 13 March 2002 (13.03.2002) US

(71) Applicant *(for all designated States except US)*: **MERCK & CO., INC.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **EMINI, Emilio, A.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **SHIVER, John, W.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **CASIMIRO, Danilo, R.** [PH/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **BETT, Andrew, J.** [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **LIANG, Xiaoping** [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **FU, Tong-Ming** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

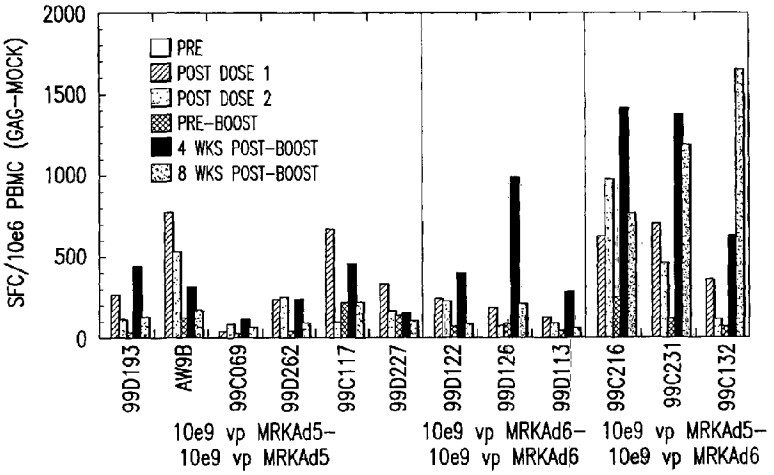
(74) Common Representative: **MERCK & CO., INC.**; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,

[Continued on next page]

(54) Title: METHOD OF INDUCING AN ENHANCED IMMUNE RESPONSE AGAINST HIV



(57) Abstract: An efficient means of inducing an immune response against human immunodeficiency virus (HIV) utilizing specific prime-boost regimes is disclosed. The specific prime-boost regimes employ a heterologous prime-boost protocol employing recombinant adenoviral vectors of alternative and distinct serotypes comprising exogenous genetic material encoding a common HIV antigen. Vaccines administered into living vertebrate tissue in accordance with the disclosed regimes, preferably a mammalian host, such as a human or a non-human mammal of commercial or domestic veterinary importance, express the HIV-1 antigen (e.g., Gag), inducing a cellular immune response which specifically recognizes HIV-1. It is believed that the disclosed prime/boost regime will offer a prophylactic advantage to previously uninfected individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-1 infection.

WO 03/077859 A2

SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

TITLE OF THE INVENTION

METHOD OF INDUCING AN ENHANCED IMMUNE RESPONSE AGAINST HIV

CROSS-REFERENCE TO RELATED APPLICATIONS

The present application claims priority to provisional application U.S. Serial No. 60/363,807, filed March 13, 2002, hereby incorporated by reference herein.

STATEMENT REGARDING FEDERALLY-SPONSORED R&D

Not Applicable

REFERENCE TO MICROFICHE APPENDIX

Not Applicable

FIELD OF THE INVENTION

The present invention relates to an enhanced means for inducing an immune response against human immunodeficiency virus ("HIV"). Recombinant adenovirus vehicles comprising exogenous genetic material encoding a common HIV antigen are employed in a heterologous prime-boost administration. More particularly, recombinant adenovirus vehicles of alternative and distinct serotypes are employed in heterologous prime-boost immunization schemes. Applicants have found that administration of a recombinant adenoviral vehicle comprising exogenous genetic material encoding an HIV antigen followed by subsequent administration of a recombinant adenovirus of a different serotype comprising the antigen notably amplifies the immune response from the initial administration(s). This amplification is, further, notably higher than that observed upon utilizing the same respective recombinant adenoviral vectors independently for both priming and boosting administrations of mammalian hosts. The amplified immune response which is particularly manifest in the cellular immune response is, further, capable of specifically recognizing HIV. Viruses of use in the instant invention can be any replication-defective adenovirus, provided that the adenovirus of choice is capable of effecting expression of exogenous genetic material incorporated into the viral sequence. Based on the findings disclosed herein, it is believed that the disclosed prime/boost regime will offer a prophylactic advantage to previously uninfected

individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-1 infection.

BACKGROUND OF THE INVENTION

Human Immunodeficiency Virus-1 (HIV-1) is the etiological agent of acquired human immune deficiency syndrome (AIDS) and related disorders. HIV-1 is an RNA virus of the Retroviridae family and exhibits the 5' LTR-*gag-pol-env*-LTR 3' organization of all retroviruses. The integrated form of HIV-1, known as the provirus, is approximately 9.8 Kb in length. Each end of the viral genome contains flanking sequences known as long terminal repeats (LTRs). The HIV genes encode at least nine proteins and are divided into three classes; the major structural proteins (Gag, Pol, and Env), the regulatory proteins (Tat and Rev); and the accessory proteins (Vpu, Vpr, Vif and Nef).

Effective treatment regimes for HIV-1 infected individuals have become available. However, these drugs will not have a significant impact on the disease in many parts of the world and they will have a minimal impact in halting the spread of infection within the human population. As is true of many other infectious diseases, a significant epidemiologic impact on the spread of HIV-1 infection will only occur subsequent to the development and introduction of an effective vaccine. There are a number of factors that have contributed to the lack of successful vaccine development to date. For instance, it is now apparent that in a chronically infected person there exists constant virus production in spite of the presence of anti-HIV-1 humoral and cellular immune responses and destruction of virally infected cells. As in the case of other infectious diseases, the outcome of disease is the result of a balance between the kinetics and the magnitude of the immune response and the pathogen replicative rate and accessibility to the immune response. Pre-existing immunity may be more successful with an acute infection than an evolving immune response can be with an established infection. A second factor is the considerable genetic variability of the virus. Although anti-HIV-1 antibodies exist that can neutralize HIV-1 infectivity in cell culture, these antibodies are generally virus isolate-specific in their activity. It has proven impossible to define serological groupings of HIV-1 using traditional methods. Rather, the virus seems to define a serological "continuum" so that individual neutralizing antibody responses, at best, are effective against only a handful of viral variants. Given this latter observation, it would be useful to identify

immunogens and related delivery technologies that are likely to elicit anti-HIV-1 cellular immune responses. It is known that in order to generate CTL responses antigen must be synthesized within or introduced into cells, subsequently processed into small peptides by the proteasome complex, and translocated into the endoplasmic reticulum/Golgi complex secretory pathway for eventual association with major histocompatibility complex (MHC) class I proteins. $CD8^+$ T lymphocytes recognize antigen in association with class I MHC via the T cell receptor (TCR) and the CD8 cell surface protein. Activation of naive $CD8^+$ T cells into activated effector or memory cells generally requires both TCR engagement of antigen as described above as well as engagement of costimulatory proteins. Optimal induction of CTL responses usually requires "help" in the form of cytokines from $CD4^+$ T lymphocytes which recognize antigen associated with MHC class II molecules via TCR and CD4 engagement.

Adenoviral vectors have been developed as live viral vectors for the delivery and expression of various foreign antigens including HIV and have proven to be effective in eliciting a significant CTL response in treated individuals. Adenoviruses are non-enveloped viruses containing a linear double-stranded genome of about 36 kb. The vectors achieve high viral titres, have a broad cell tropism, and can infect nondividing cells. Adenoviral vectors are very efficient gene transfer vehicles and are frequently used in clinical gene therapy studies. In addition, adenovirus has formed the basis of many promising viral immunization protocols.

European Patent Applications 0 638 316 (Published February 15, 1995) and 0 586 076 (Published March 9, 1994), (both assigned to American Home Products Corporation) describe replicating adenovirus vectors carrying an HIV gene, including *env* or *gag*. Various treatment regimes based on these vectors were used with chimpanzees and dogs, some of which included booster adenovirus or protein plus alum treatments.

Replication-defective adenoviral vectors harboring deletions, for instance, in the E1 region constitute a safer alternative to their replicating counterparts. Recent adenoviral vectors have incorporated the known packaging repeats into these vectors; e.g., see EP 0 707 071, disclosing, *inter alia*, an adenoviral vector deleted of E1 sequences from base pairs 459 to 3328; and U.S. Patent No. 6,033,908, disclosing, *inter alia*, an adenoviral vector deleted of base pairs 459-3510. The packaging efficiency of adenovirus has been taught to depend on the number of incorporated

individual A (packaging) repeats; *see, e.g.,* Gräble and Hearing, 1990 *J. Virol.* 64(5):2047-2056; Gräble and Hearing, 1992 *J. Virol.* 66(2):723-731.

Adenovirus serotypes 5 and 6 have been disclosed and are publicly available (*see,* American Type Culture Collection ("ATCC") Accession Deposit Nos. VR-5 and VR-6; respectively). The wildtype adenovirus serotype 5 sequence is, further, known and described in the art; *see*, Chroboczek *et al.*, 1992 *J. Virology* 186:280-5. The complete sequence for adenovirus serotype 6, which is provided in Figures 11A-1 to 11A-14, was first disclosed in copending U.S. Provisional Application Serial No. 60/328,655, filed on October 11, 2001. Adenovirus serotype 6, as serotype 5, has been described previously in the literature; *see* Rowe *et al.*, 1953 *Proc. Soc. Exp. Biol. Med.* 84:570; Rowe *et al.*, 1955 *Am. J. Hyg.* 61:197-218; and Hierholzer *et al.*, 1991 *Arch. Virol.* 121:179-97. Adenovirus serotypes other than Ad5 and Ad6 are also known and described in the literature.

Administration protocols employing viral vaccine vectors to date have employed various prime-boost inoculation schemes. Two general schemes frequently used are: (1) wherein both priming and boosting of the mammalian host is accomplished using the same virus vehicle, and (2) wherein the priming and boosting is carried out utilizing different vehicles not necessarily limited to virus vehicles. Examples of the latter are, for instance, a scheme composed of a DNA prime and viral boost, and one composed of a viral prime and a viral boost wherein alternate virus are used.

It would be of great import in the battle against AIDS to develop a prophylactic- and/or therapeutic-based HIV vaccine strategy capable of generating a strong cellular immune response against HIV infection. The present invention addresses and meets these needs by disclosing a heterologous prime-boost HIV immunization regime based on the administration of recombinant adenoviral vectors of alternative and distinct serotypes, wherein the recombinant adenoviral vectors comprise exogenous genetic material encoding a common HIV antigen. One aspect of the instant invention concerns heterologous immunization schemes employing recombinant adenoviral vectors derived from adenovirus serotypes 5, 6, and 35. A vaccine protocol in accords with this description, as far as Applicants are aware, has not been demonstrated for HIV. This vaccine prime-boost regime may be administered to a host, such as a human.

SUMMARY OF THE INVENTION

The present invention relates to an enhanced method for generating an immune response against human immunodeficiency virus ("HIV"). The method is based on the heterologous prime-boost administration of recombinant adenovirus vehicles of alternative and distinct serotypes comprising heterologous genetic material encoding an HIV antigen to effect a more pronounced immune response against HIV than that which can be obtained by either vector independently in a single modality prime-boost immunization scheme. In accordance with the disclosed methods, a mammalian host is first administered a priming dose comprising a recombinant adenoviral vector of a first serotype comprising a gene encoding an HIV antigen and, after a period of time, administered a boosting dose comprising a recombinant adenoviral vector of a second and different serotype carrying the gene encoding the HIV antigen. There may be a predetermined minimum amount of time separating the administrations, which time essentially allows for an immunological rest. In particular embodiments, this rest is for a period of at least 4 months. Multiple primings typically, 1-4, are usually employed, although more may be used. The length of time between priming and boost may typically vary from about four months to a year, but other time frames may be used. Applicants have found that boosting of the adenovirus-primed response with an adenovirus of an alternative and distinct serotype leads to a notably amplified immune response to the HIV antigen. Thus the instant invention relates to the administration of alternate serotype adenovirus HIV vaccines in accordance with the disclosed methods.

Accordingly, the instant invention relates to a method for inducing an enhanced immunological response against an HIV-1 antigen in a mammalian host comprising the steps of (a) inoculating the mammalian host with a recombinant adenoviral vector of a first serotype which is at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding an HIV-1 antigen or an immunologically relevant modification thereof; and thereafter (b) inoculating the mammalian host with a boosting immunization comprising a recombinant adenoviral vector of a second and different serotype at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding an HIV-1 antigen or immunologically relevant modification thereof.

The recombinant adenoviral vectors used in the immunization regimes of the present invention may comprise any replication-defective adenoviral vector which is

genetically stable through large-scale production and purification of the virus. In other words, a recombinant adenoviral vector suitable for use in the methods of the instant invention can be any purified recombinant replication-defective virus shown to be genetically stable through multiple passages in cell culture which remains so during large-scale production and purification procedures. Such a recombinant virus vector and harvested virus vaccine lends itself to large scale dose filling and subsequent worldwide distribution procedures which will be demanded of an efficacious monovalent or multivalent HIV vaccine. The present invention meets this basic requirement with description of an immunization regime which is based on the use of recombinant replication-defective adenovirus serotypes examples but not limitations of which include serotypes 5, 6, and 35.

Adenoviral vectors preferred for use in the immunization regimes of the instant invention are those that are at least partially deleted in E1 and devoid of E1 activity. Vectors in accordance with this description can be readily propagated in E1-complementing cell lines, such as PER.C6® cells.

The recombinant adenoviral vectors of use in the instant application whether intended as the priming or boosting vehicle must comprise a gene encoding an HIV antigen. In specific embodiments, the gene encoding the HIV antigen or immunologically relevant modification thereof comprises codons optimized for expression in a mammalian host (*e.g.*, a human). Recombinant adenoviral vectors of use in the methods of the instant invention can comprise a gene expression cassette comprising (a) nucleic acid encoding an HIV antigen (*e.g.*, an HIV protein) or biologically active and/or immunologically relevant portion thereof; (b) a heterologous (non-native) or modified native promoter operatively linked to the nucleic acid of part a); and, (c) a transcription termination sequence. A heterologous promoter can be any promoter under the sun (modified or not) which is not native to, or derived from, the virus in which it will be used.

HIV antigens of use in the instant invention include the various HIV proteins, immunologically relevant modifications, and immunogenic portions thereof. The present invention, thus, encompasses the various forms of codon-optimized HIV-1 gag (including but by no means limited to p55 versions of codon-optimized full length ("FL") Gag and tPA-Gag fusion proteins), HIV-1 pol, HIV-1 nef, HIV-1 env, fusions of the above constructs, and selected modifications of the above possessing immunological relevance. Examples of HIV-1 Gag, Pol, Env, and/or Nef fusion

proteins include but are not limited to fusion of a leader or signal peptide at the NH_2-teriminal portion of the viral antigen coding region. Such a leader peptide includes but is not limited to a tPA leader peptide.

Recombinant viral vectors in accordance with the instant disclosure form an aspect of the instant invention. Other aspects of the instant invention are host cells comprising said adenoviral vectors; vaccine compositions comprising said vectors; and methods of producing the vectors comprising (a) introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and (b) harvesting the resultant adenoviral vectors.

The present invention also relates to prime-boost regimes wherein the recombinant adenoviral vectors comprise various combination of the above HIV antigens. Such HIV immunization regimes will provide for an enhanced cellular immune response subsequent to host administration, particularly given the genetic diversity of human MHCs and of circulating virus. Examples, but not limitations, include viral vector-based multivalent vaccine compositions which provide for a divalent (*e.g.*, gag and nef, gag and pol, or pol and nef components) or a trivalent vaccine (*e.g.*, gag, pol and nef components) composition. Such a multivalent vaccine may be filled for a single dose or may consist of multiple inoculations of each individually filled component. To this end, preferred vaccine compositions of use in the methods of the instant application are recombinant adenovirus vectors comprising multiple, distinct HIV antigen classes. Each HIV antigen class is subject to sequence manipulation, thus providing for a multitude of potential vaccine combinations; and such combinations are within the scope of the present invention. The utilization of such combined modalities increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a single modality regime.

The concept of a "combined modality" as disclosed herein also covers the alternative mode of administration whereby multiple HIV-1 viral antigens may be ligated into a proper shuttle plasmid for generation of a recombinant viral vector comprising multiple open reading frames. For example, a trivalent vector may comprise a gag-pol-nef fusion, or possibly a "2+1" divalent vaccine comprising, for instance, a gag-pol fusion (i.e., codon optimized p55 gag and inactivated optimized pol) within the same backbone, with each open reading frame being operatively linked to a distinct promoter and transcription termination sequence. Alternatively, the

two open reading frames may be operatively linked to a single promoter, with the open reading frames operatively linked by an internal ribosome entry sequence (IRES).

Administration of the recombinant adenoviral vectors via the disclosed heterologous means provides for improved cellular-mediated immune responses; responses more pronounced than that afforded by single modality regimes. An effect of the improved vaccine should be a lower transmission rate to previously uninfected individuals (i.e., prophylactic applications) and/or reduction in the levels of the viral loads within an infected individual (i.e., therapeutic applications), so as to prolong the asymptomatic phase of HIV-1 infection. The administration, intracellular delivery and expression of the vaccine in this manner elicits a host CTL and Th response. The individual vaccinee or mammalian host (as referred to herein) can be a primate (both human and non-human) as well as any non-human mammal of commercial or domestic veterinary importance.

In light hereof, the present invention relates to methodology regarding administration of the recombinant adenoviral HIV vaccines to provide effective immunoprophylaxis, to prevent establishment of an HIV-1 infection following exposure to this virus, or as a post-HIV infection therapeutic vaccine to mitigate the acute HIV-1 infection so as to result in the establishment of a lower virus load with beneficial long term consequences. Such treatment regimes may include a monovalent or multivalent composition, and/or various combined modality applications. Therefore, the present invention provides for methods of using the disclosed HIV vaccine administration scheme within the various parameters disclosed herein as well as any additional parameters known in the art which, upon introduction into mammalian tissue, induces intracellular expression of the HIV antigen(s) and an effective immune response to the respective HIV antigen(s).

To this end, the present invention relates in part to methods of generating a cellular immune response in a vaccinee, preferably a human vaccinee, wherein the individual is given the recombinant adenovirus HIV vaccines in the manner described.

As used throughout the specification and claims, the following definitions and abbreviations are used:

"HAART" refers to -- highly active antiretroviral therapy --.

"first generation" vectors are characterized as being replication-defective. They typically have a deleted or inactivated E1 gene region, and often have a deleted or inactivated E3 gene region as well.

"AEX" refers to Anion Exchange chromatography.

"QPA" refers to Quick PCR-based Potency Assay.

"bps" refers to base pairs.

"s" or "str" denotes that the transgene is in the E1 parallel or "straight" orientation.

"PBMCs" refers to peripheral blood monocyte cells.

"FL" refers to full length.

"FLgag" refers to a full-length optimized gag gene, as shown in Figure 2.

"Ad5-Flgag" refers to an adenovirus serotype 5 replication-deficient virus which carries an expression cassette which comprises a full length optimized gag gene under the control of a CMV promoter.

"Promoter" means a recognition site on a DNA strand to which an RNA polymerase binds. The promoter forms an initiation complex with RNA polymerase to initiate and drive transcriptional activity. The complex can be modified by activating sequences such as enhancers or inhibiting sequences such as silencers.

"Leader" means a DNA sequence at the 5' end of a structural gene which is transcribed along with the gene. This usually results in a protein having an N-terminal peptide extension, often referred to as a pro-sequence.

"Intron" means a section of DNA occurring in the middle of a gene which does not code for an amino acid in the gene product. The precursor RNA of the intron is excised and therefore not transcribed into mRNA or translated into protein.

"Immunologically relevant" or "biologically active," when used in the context of a viral protein, means that the protein is capable, upon administration, of eliciting a measurable immune response within an individual sufficient to retard the propagation and/or spread of the virus and/or to reduce the viral load present within the individual. The same terms, when used in the context of a nucleotide sequence, means that the sequence is capable of encoding for a protein capable of the above.

"Cassette" refers to a nucleic acid sequence which is to be expressed, along with its transcription and translational control sequences. By changing the cassette, a vector can express a different sequence.

"bGHpA" refers to a bovine growth hormone transcription terminator/polyadenylation sequence.

"tPAgag" refers to a fusion between the tissue plasminogen activator leader sequence and an optimized HIV gag gene.

Where utilized, "IA" or "inact" refers to an inactivated version of a gene (e.g. IApol).

"MCS" is "multiple cloning site".

"Ad5" is adenovirus of serotype 5.

"Ad6" is adenovirus of serotype 6.

In general, adenoviral constructs, gene constructs are named by reference to the genes contained therein. For example:

"Ad5 HIV-1 gag", also referred to as the original HIV-1 gag adenoviral vector, is a vector containing a transgene cassette composed of a hCMV intron A promoter, the full length version of the human codon-optimized HIV-1 gag gene, and the bovine growth hormone polyadenylation signal.

"MRK Ad5 HIV-1 gag" also referred to as "MRKAd5gag" or "Ad5gag2" is an adenoviral vector which is deleted of E1, and contains adenoviral base pairs 1-450 and 3511-3523, with a human codon-optimized HIV-1 gag gene in an E1 parallel orientation under the control of a CMV promoter without intron A. The construct also comprises a bovine growth hormone polyadenylation signal.

"pV1JnsHIVgag", also referred to as "HIVFLgagPR9901", is a plasmid comprising the CMV immediate-early (IE) promoter and intron A, a full-length codon-optimized HIV gag gene, a bovine growth hormone-derived polyadenylation and transcriptional termination sequence, and a minimal pUC backbone.

"pV1JnsCMV(no intron)-FLgag-bGHpA" is a plasmid derived from pV1JnsHIVgag which is deleted of the intron A portion of CMV and which comprises the full length HIV gag gene. This plasmid is also referred to as "pV1JnsHIVgag-bGHpA", pV1Jns-hCMV-FL-gag-bGHpA" and "pV1JnsCMV(no intron) + FLgag + bGHpA".

"pV1JnsCMV(no intron)-FLgag-SPA" is a plasmid of the same composition as pV1JnsCMV(no intron)-FLgag-bGHpA except that the SPA termination sequence replaces that of bGHpA. This plasmid is also referred to as "pV1Jns-HIVgag-SPA" and pV1Jns-hCMV-FLgag-SPA".

"pdelE1sp1A" is a universal shuttle vector with no expression cassette (i.e., no promoter or polyA). The vector comprises wildtype adenovirus serotype 5 (Ad5) sequences from bp 1 to bp 341 and bp 3524 to bp 5798, and has a multiple cloning site between the Ad5 sequences ending 341 bp and beginning 3524 bp. This plasmid is also referred to as the original Ad 5 shuttle vector.

"MRKpdelE1sp1A" or "MRKpdelE1(Pac/pIX/pack450)" or "MRKpdelE1(Pac/pIX/pack450)Cla1" is a universal shuttle vector with no expression cassette (i.e. no promoter or polyA) comprising wildtype adenovirus serotype 5 (Ad5) sequences from bp 1 to bp 450 and bp 3511 to bp 5798. The vector has a multiple cloning site between the Ad5 sequence ending 450 bp and beginning 3511 bp. This shuttle vector may be used to insert the CMV promoter and the bGHpA fragments in both the straight ("str". or E1 parallel) orientation or in the opposite (opp. or E1 antiparallel) orientation.

"MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.)" is still another shuttle vector which is the modified vector that contains the CMV promoter (no intron A) and the bGHpA fragments. The expression unit containing the hCMV promoter (no intron A) and the bovine growth hormone polyadenylation signal has been inserted into the shuttle vector such that insertion of the gene of choice at a unique *Bgl*II site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1/E3+)Cla1 pre-plasmid.

"MRKpdelE1-CMV(no intron)-FLgag-bGHpA" is a shuttle comprising Ad5 sequences from base pairs 1-450 and 3511-5798, with an expression cassette containing human CMV without intron A, the full-length human codon-optimized HIV gag gene and bovine growth hormone polyadenylation signal. This plasmid is also referred to as "MRKpdelE1 shuttle +hCMV-FL-gag-BGHpA"

"MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA" is an adenoviral vector comprising all Ad5 sequences except those nucleotides encompassing the E1 region (from 451-3510), a human CMV promoter without intron A, a full-length human codon-optimized HIV gag gene, and a bovine growth hormone polyadenylation signal. This vector is also referred to as "MRKpAdHVE3 + hCMV-FL-gag-BGHpA", "MRKpAd5HIV-1gag", "MRKpAd5gag", "pMRKAd5gag" or "pAd5gag2".

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the HIV-1 gag adenovector "Ad5 HIV-1 gag". This vector is disclosed in PCT International Application No. PCT/US00/18332 (WO 01/02607) filed July 3, 2000, claiming priority to U.S. Provisional Application Serial No. 60/142,631, filed July 6, 1999, and U.S. Application Serial No. 60/148,981, filed August 13, 1999, all three applications which are hereby incorporated by reference.

Figure 2 shows the nucleic acid sequence (SEQ ID NO: 1) of the optimized human HIV-1 gag open reading frame.

Figure 3 shows diagrammatically the transgene construct disclosed in PCT International Application No. PCT/US01/28861, filed September 14, 2001 in comparison with the original gag transgene. PCT International Application No. PCT/US01/28861 claims priority to U.S. Provisional Application Serial Nos. 60/233,180, 60/279,056, and 60/317,814, filed September 15, 2000, March 27, 2001, and September 7, 2001, respectively; the above applications all of which are hereby incorporated by reference.

Figure 4 shows the modifications made to the adenovector backbone of Ad5HIV-1gag in the generation of the vector disclosed in PCT International Application No. PCT/US01/28861 which is utilized in certain examples of the instant application.

Figure 5 shows the levels of Gag-specific T cells in rhesus macaques immunized with (a) two priming doses of 10e9 vp of MRKAd5 HIV-1 gag and a single booster shot with 10e9 vp MRKAd5 HIV-1 gag ("10e9 vp MRKAd5-10e9 vp MRKAd5"); (b) two priming doses of 10e9 pfu MRKAd6 HIV-1 gag and a single booster with 10e9 pfu MRKAd6 HIV-1 gag ("10e9 pfu MRKAd6-10e9 pfu MRKAd6"); or (c) two priming doses of 10e9 vp of MRKAd5 HIV-1 gag followed by a single booster shot with 10e9 pfu MRKAd6 HIV-1 gag ("10e9 vp MRKAd5-10e9 pfu MRKAd6"). The levels expressed as number of spot-forming cells (SFC) per million PBMC are the mock-corrected values for each animal prior to the start of the immunization regimen ("Pre"); 4 weeks after the first priming dose ("Post Dose 1"); 4 weeks after the second priming dose ("Post Dose 2"); just prior to the boost ("Pre-Boost"); 4 weeks after the boost ("4 wks Post-Boost"); and 8 weeks after the boost ("8 wks Post-Boost").

Figure 6 shows the Gag-specific T cell responses induced by two priming doses of 10e7 vp dose of MRKAd5 HIV-1 gag (week 0; week 4) followed by

administration of 10e7 vp MRKAd6 HIV-1 gag at week 27. The levels provided are the mock-corrected levels for each animal prior to the start of the immunization regimen ("Pre"); 4 weeks after the first priming dose ("Post Dose 1"); 4 weeks after the second priming dose ("Post Dose 2"); just prior to the boost ("Pre-Boost"); 4 weeks after the boost ("4 wks Post-Boost"); and 8 weeks after the boost ("8wks Post-Boost"). One will note a significant increase compared to the levels just prior to the boost. MRKAd6 HIV-1 gag elicited a large amplification of the priming response. The post-boost increases shown are largely attributed to the expansion of memory T cells instead of priming of new lymphocytes.

Figure 7 shows the homologous recombination protocol utilized to recover pAdE1-E3 disclosed herein.

Figure 8 shows a restriction map of the pMRKAd5HIV-1gag vector.

Figures 9A-1 to 9A-45 show the nucleotide sequence of the pMRKAd5HIV-1gag vector (SEQ ID NO:2 [coding] and SEQ ID NO:3 [non-coding]).

Figure 10 shows the levels of Gag-specific antibodies in rhesus macaques immunized with (a) two priming doses of 10e9 vp of MRKAd5 HIV-1 gag and a single booster shot with 10e9 vp MRKAd5 HIV-1 gag ("10e9 vp MRKAd5-10e9 vp MRKAd5"), (b) two priming doses of 10e9 pfu MRKAd6 HIV-1 gag and a single booster with 10e9 pfu MRKAd6 HIV-1 gag ("10e9 pfu MRKAd6-10e9 pfu MRKAd6"), or (c) two priming doses of 10e9 vp of MRKAd5 HIV-1 gag followed by a single booster shot with 10e9 pfu MRKAd6 HIV-1 gag ("10e9 vp MRKAd5-10e9 pfu MRKAd6"). Shown are the geometric mean titers for each cohort at the start of the immunization regimen ("Pre"), 4 weeks after the first priming dose ("Wk 4"), 4 weeks after the second priming dose ("Wk 8"), just prior to the boost ("Pre-Boost"), and 8 weeks after the boost ("Post-Boost").

Figures 11A-1 to 11A-14 show the nucleic acid sequence for the Ad6 genome (SEQ ID NO:5).

Figure 12 shows the basic genomic organization of Ad6. The linear (35759 bp) double-stranded DNA genome is indicated by two parallel lines and is divided into 100 map units. Transcription units are shown relative to their position and orientation in the genome. Early genes (E1A, E1B, E2A/B, E3 and E4) are indicated by gray bars. Late genes (L1 to L5), indicated by black bars, are produced by alternative splicing of a transcript produced from the major late promoter (MLP) and all contain the tripartite leader (1, 2, 3) at their 5' ends.

Figure 13 shows the homologous recombination protocol utilized to recover pMRKAd6E1-.

DETAILED DESCRIPTION OF THE INVENTION

An enhanced means for generating an immune response against human immunodeficiency virus ("HIV") is described. The disclosed methods employ a combination of recombinant adenovirus gene delivery vehicles of alternative and distinct serotypes in the administration of exogenous genetic material encoding an HIV antigen (or antigens) of interest. In accordance with the methods of the instant invention, a priming dose of the HIV antigen(s) is first delivered with a recombinant adenoviral vector of a first serotype. This dose effectively primes the immune response so that, upon subsequent identification of the antigen in the circulating immune system, the immune response is capable of immediately recognizing and responding to the antigen within the host. The priming dose(s) is then followed up with a boosting dose of a second and different adenovirus serotype comprising exogenous genetic material encoding the antigen. In one aspect of the instant invention, a mammalian host is first administered a priming dose(s) comprising a recombinant adenoviral vector of serotype 5 or 6 and then administered a subsequent boosting dose(s) comprising a recombinant adenoviral vector of a different serotype (*i.e.,* a serotype other than that used in the priming administration; examples, but not limitations of which include Ad35. Very specific embodiments encompassed herein are wherein (1) an Ad5-primed response is boosted with a recombinant Ad6 vehicle comprising an HIV antigen; (2) an Ad6-primed response is boosted with a recombinant Ad5 vehicle comprising an HIV antigen; (3) an Ad5/Ad6-primed response is boosted with a recombinant, Ad35-based vehicle; and (4) an Ad35-primed response is boosted with a recombinant, an Ad5/Ad6-based vehicle. As relates to HIV antigens, administration in accordance with the methods of the instant invention results in a significant non-additive synergistic effect which notably increases the immune response seen in inoculated mammalian hosts. The effects are particularly evident in the cellular immune responses generated following inoculation. The disclosed immunization regime, thus, offers a prophylactic advantage to previously uninfected individuals and can offer a therapeutic effect to reduce viral load levels in those already infected with the virus, thus prolonging the asymptomatic phase of HIV-1 infection.

Accordingly, the instant invention relates to a method for inducing an enhanced immunological response against an HIV-1 antigen in a mammalian host comprising the steps of (a) inoculating the mammalian host with a recombinant adenoviral vector of a first serotype at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding an HIV-1 antigen or immunologically relevant modification thereof; and thereafter (b) inoculating the mammalian host with a boosting immunization comprising a recombinant adenovirus vector of a second and distinct serotype at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding an HIV-1 antigen or immunologically relevant modification thereof. Preferred embodiments of the instant invention employ adenoviral vectors which are replication-defective by reason of having a deletion in the E1 region which renders the vector devoid (or essentially devoid) of E1 activity. Adenovirus serotype 5 has been found to be a very effective adenovirus vehicle for purposes of effectuating sufficient expression of exogenous genetic material encoding HIV-specific antigens in order to provide for sufficient priming of the mammalian host immune response. It has further been found and disclosed herein that recombinant adenovirus serotype 6 is capable of very effectively boosting the adenovirus serotype 5-primed response. In an alternative scenario, recombinant adenovirus serotype 5 can be used to boost an adenovirus serotype 6-primed response. These findings have also been demonstrated with adenovirus vehicles of different subgroups, for instance, Ad5/6-prime (subgroup C)/Ad35-boost (subgroup B).

The wildtype adenovirus serotype 5 sequence is known and described in the art; *see*, Chroboczek *et al.*, 1992 *J. Virology* 186:280, which is hereby incorporated by reference. Accordingly, a particular embodiment of the instant invention is an immunization scheme employing an adenovirus vehicle based on the wildtype adenovirus serotype 5 sequence in the priming or boosting administration; a virus of which is on deposit with the American Type Culture Collection ("ATCC") under ATCC Deposit No. VR-5. One of skill in the art can, however, readily identify alternative and distinct adenovirus serotypes (e.g., serotypes 2, 4, 6, 12, 16, 17, 24, 31, 33, and 42) and incorporate same in the disclosed heterologous prime-boost immunization schemes. The sequence of adenovirus serotype 6 (ATCC Deposit No. VR-6) is extremely homologous (approximately 98%) at the nucleic acid level to the sequence of adenovirus serotype 5, with relatively few base pair differences in the approximate 36 kb sequences. The genomic organization of Ad6 is also very similar;

see Figure 12. Chimeric Ad5/Ad6 constructs which retain the serotype-determining epitopes of either Ad5 or Ad6 are also suitable for use in the instant invention; provided that the serotype determining epitopes are distinct from the adenovirus vehicle used in combination therewith (*i.e.,* that the determinants are distinct from the vehicle used in the priming dose if the chimera is utilized in the boosting dose, and *vice versa*). It is important to the overall functioning of the disclosed methods that the serotypes of the priming and boosting vectors be distinct.

Recombinant adenoviral vectors comprising deletions additional to that contained within the region of E1 are also contemplated for use within the methods of the instant invention. For example, vectors comprising deletions in both E1 and E3 are contemplated for use within the methods of the instant invention. Such a vector can accommodate a larger amount of foreign DNA (or exogenous genetic material).

Adenoviral vectors of use in the methods of the instant invention can be constructed using known techniques, such as those reviewed in Hitt *et al.*, 1997 "Human Adenovirus Vectors for Gene Transfer into Mammalian Cells" *Advances in Pharmacology* 40:137-206, which is hereby incorporated by reference. Often, a plasmid or shuttle vector is generated which comprises sequence from the specific adenovirus of interest. This process is described in Hitt *et al., supra.*

Adenoviral pre-plasmids (e.g., pMRKAd5gag and pMRKAd6gag) can be generated by homologous recombination using adenovirus backbones (*e.g.,* MRKAd5HVE3 and pMRKAd6E1-, an Ad6 genome plasmid) and the appropriate shuttle vector. The resultant plasmids in linear form, are capable of replication after entering the PER.C6® cells or other complementing cell line, and virus is produced. The infected cells and media are then harvested after viral replication is complete.

Viral vectors can be propagated in various E1 complementing cell lines, including the known cell lines 293 and PER.C6®. Both these cell lines express the adenoviral E1 gene product. PER.C6® is described in WO 97/00326 (published January 3, 1997) and issued U.S. Patent No. 6,033,908, both of which are hereby incorporated by reference. It is a primary human retinoblast cell line transduced with an E1 gene segment that complements the production of replication deficient (FG) adenovirus, but is designed to prevent generation of replication competent adenovirus by homologous recombination. Cells of particular interest have been stably transformed with a transgene that encodes the AD5E1A and E1B gene, like PER.C6®, from 459 bp to 3510 bp inclusive. 293 cells are described in Graham et al., 1977 *J.*

Gen. Virol 36:59-72, which is hereby incorporated by reference. As stated above, due consideration must be given to the adenoviral sequences present in the complementing cell line used. It is preferred that the sequences not overlap with that present in the vector if the possibility of recombination is to be minimized.

The recombinant adenoviral vectors of use in the instant invention comprise a gene encoding any antigen, but particularly, an HIV-1 antigen or an immunologically relevant modification thereof. HIV antigens of interest include, but are not limited to, the major structural proteins of HIV such as Gag, Pol, and Env, immunologically relevant modifications, and immunogenic portions thereof. The invention, thus, encompasses the various forms of codon-optimized HIV-1 gag (including but by no means limited to p55 versions of codon-optimized full length ("FL") Gag and tPA-Gag fusion proteins), HIV-1 pol, HIV-1 nef, HIV-1 env, and selected modifications of immunological relevance.

Exogenous genetic material encoding a protein of interest may exist in the form of an expression cassette. A gene expression cassette preferably comprises (a) a nucleic acid encoding a protein of interest; (b) a heterologous (non-native) or modified native promoter operatively linked to the nucleic acid encoding the protein; and (c) a transcription termination sequence.

The transcriptional promoter is preferably recognized by an eukaryotic RNA polymerase. In a preferred embodiment, the promoter is a "strong" or "efficient" promoter. An example of a strong promoter is the immediate early human cytomegalovirus promoter (Chapman et al, 1991 *Nucl. Acids Res.* 19:3979-3986, which is incorporated by reference); in certain embodiments without intronic sequences. Specific embodiments of the instant invention employ human CMV promoters without intronic sequences, like intron A. Applicants have found that intron A, a portion of the human cytomegalovirus promoter (hCMV), constitutes a region of instability for adenoviral vectors. CMV without intron A has been found to effectuate comparable expression capabilities *in vitro* when driving HIV gag expression and, furthermore, behaved equivalently to intron A-containing constructs in Balb/c mice *in vivo* with respect to their antibody and T-cell responses at both dosages of plasmid DNA tested (20 μg and 200 μg). Those skilled in the art will appreciate that any of a number of other known promoters, such as the strong immunoglobulin, or other eukaryotic gene promoters may also be used, including the EF1 alpha promoter, the murine CMV promoter, Rous sarcoma virus (RSV)

promoter, SV40 early/late promoters and the beta-actin promoter. In certain embodiments, the promoter may also comprise a regulatable sequence such as the Tet operator sequence. This would be extremely useful, for example, in cases where the gene products are effecting a result other than that desired and repression is sought. Preferred transcription termination sequences present within the gene expression cassette are the bovine growth hormone terminator/polyadenylation signal (bGHpA) and the short synthetic polyA signal (SPA) of 50 nucleotides in length, defined as follows: AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGT-TTTTTGTGTG (SEQ ID NO:4). The combination of the CMV promoter (devoid of the intron A region) with the BGH terminator constitutes a specific embodiment of the present invention, although other promoter/terminator combinations can be used. Certain embodiments may incorporate a leader or signal peptide into the transgene. A preferred leader is that from the tissue-specific plasminogen activator protein, tPA.

In accordance with the methods of the instant invention, the expression of exogenous HIV genetic material should elicit potent and broad cellular immune responses against HIV that will either lessen the likelihood of persistent virus infection and/or lead to the establishment of a clinically significant lowered virus load subject to HIV infection or in combination with HAART therapy, mitigate the effects of previously established HIV infection (antiviral immunotherapy(ARI)). While any HIV antigen (e.g., gag, pol, nef, gp160, gp41, gp120, tat, rev, etc.) may be incorporated into the recombinant adenoviral vectors of use in the instant invention, preferred embodiments include the codon optimized p55 gag antigen, pol and nef. The adenoviral vehicles of the instant invention can utilize heterologous nucleic acid which may or may not be codon-optimized. In specific embodiments of the instant invention, the individual can be primed with an adenoviral vector comprising codon-optimized heterologous nucleic acid, and boosted with an adenovirus of an alternative serotype comprising non-codon-optimized nucleic acid. Administration of multiple antigens possesses the possibility for exploiting various different combinations of codon-optimized and non-codon-optimized sequences.

Sequences based on different Clades of HIV-1 are suitable for use in the instant invention, most preferred of which are Clade B and Clade C. Particularly preferred embodiments are those sequences (especially, codon-optimized sequences) based on consensus Clade B sequences. Preferred versions of the viral vaccines will encode modified versions of pol or nef. Preferred embodiments of the viral vaccines

carrying HIV envelope genes and modifications thereof comprise the HIV codon-optimized *env* sequences of PCT International Applications PCT/US97/02294 and PCT/US97/10517, published August 28, 1997 (WO 97/31115) and December 24, 1997, respectively; both documents of which are hereby incorporated by reference.

Sequences for many genes of many HIV strains are publicly available in GENBANK and primary, field isolates of HIV are available from the National Institute of Allergy and Infectious Diseases (NIAID) which has contracted with Quality Biological (Gaithersburg, MD) to make these strains available. Strains are also available from the World Health Organization (WHO), Geneva Switzerland. It is preferred that the gag gene be from an HIV-1 strain (CAM-1; Myers et al, eds. "Human Retroviruses and AIDS: 1995, IIA3-IIA19, which is hereby incorporated by reference). This gene closely resembles the consensus amino acid sequence for the clade B (North American/European) sequence. Therefore, it is within the purview of the skilled artisan to choose an appropriate nucleotide sequence which encodes a specific HIV gag antigen, or immunologically relevant portion thereof. A clade B or clade C based p55 gag antigen will potentially be useful on a global scale. A transgene of choice for insertion into the vectors utilized within the disclosed methods is a codon-optimized version of p55 gag.

In addition to a single HIV antigen of interest being delivered by the recombinant adenoviral vectors, two or more antigens can be delivered either via separate vehicles or delivered *via* the same vehicle. For instance, a priming dose in accordance with the instant invention can comprise a recombinant adenoviral vector of a first serotype comprising genes encoding both nef and pol or, alternatively, two or more alternative HIV-1 antigens. The boosting dose could then comprise a recombinant adenoviral vector of a second and different serotype comprising the genes encoding both nef and pol (or whichever two or more HIV-1 antigens were used in the priming dose). In an alternative scenario, the priming dose can comprise a mixture of separate adenoviral vehicles each comprising a gene encoding for a different HIV-1 antigen. In such a case, the boosting dose could also comprise a mixture of vectors each comprising a gene encoding for a separate HIV-1 antigen, provided that the boosting dose(s) administers recombinant viral vectors comprising genetic material encoding for the same or a similar set of antigens that were delivered in the priming dose(s). These divalent (*e.g.*, gag and nef, gag and pol, or pol and nef components, for instance) or trivalent (*e.g.*, gag, pol and nef components, for instance)

vaccines can further be administered by a combination of the techniques described above. Therefore, a preferred aspect of the present invention are the various vaccine formulations that can be administered by the methods of the instant invention. It is also within the scope of the present invention to embark on combined modality regimes which include multiple but distinct components from a specific antigen.

The disclosed immunization regimes employing fusion constructs composed of two or more antigens are also encompassed herein. For example, multiple HIV-1 viral antigens may be ligated into a proper shuttle plasmid for generation of a pre-viral plasmid comprising multiple open reading frames. For example a trivalent vector may comprise a gag-pol-nef fusion, or possible a "2+1" divalent vaccine comprising, for instance, a gag-pol fusion (*e.g.*,, a codon optimized p55 gag and inactivated optimized pol) with each open reading frame being operatively linked to a distinct promoter and transcription termination sequence. Alternatively, the two open reading frames in the same construct may be operatively linked to a single promoter, with the open reading frames operatively linked by an internal ribosome entry sequence (IRES), as disclosed in International Publication No. WO 95/24485, which is hereby incorporated by reference. In the absence of the use of IRES-based technology, it is preferred that a distinct promoter be used to support each respective open reading frame, so as to best preserve vector stability. As examples, and certainly not as limitations, potential multiple transgene vaccines may include a three transgene vector such as that wherein a gagpol fusion and nef gene were included in the same vector with different promoters and termination sequences being used for the gagpol fusion and nef gene. Further, potential "2+1" divalent vaccines of the present invention might be wherein a construct containing gag and nef in the same construct with separate promoters and termination sequences is administered in combination with a construct comprising a pol gene with promoter and termination sequence. Fusion constructs other than the gag-pol fusion described above are also suitable for use in various divalent vaccine strategies and can be composed of any two HIV antigens fused to one another (*e.g.*,, nef-pol and gag-nef). These compositions are, as above, preferably delivered along with a viral composition comprising an additional HIV antigen in order to diversify the immune response generated upon inoculation. Therefore, a multivalent vaccine delivered in a single, or possibly second, viral vector is certainly contemplated as part of the present invention. It is important to note, however, that in terms of deciding on an insert for the disclosed viral vectors, due

consideration must be given to the effective packaging limitations of the viral vehicle. Adenovirus, for instance, has been shown to exhibit an upper cloning capacity limit of approximately 105% of the wildtype Ad5 sequence.

Regardless of the gene chosen for expression, it is preferred that the sequence be "optimized" for expression in a mammalian (e.g., human) cellular environment, particularly in the adenoviral constructs. A "triplet" codon of four possible nucleotide bases can exist in 64 variant forms. That these forms provide the message for only 20 different amino acids (as well as transcription initiation and termination) means that some amino acids can be coded for by more than one codon. Indeed, some amino acids have as many as six "redundant", alternative codons while some others have a single, required codon. For reasons not completely understood, alternative codons are not at all uniformly present in the endogenous DNA of differing types of cells and there appears to exist variable natural hierarchy or "preference" for certain codons in certain types of cells. As one example, the amino acid leucine is specified by any of six DNA codons including CTA, CTC, CTG, CTT, TTA, and TTG (which correspond, respectively, to the mRNA codons, CUA, CUC, CUG, CUU, UUA and UUG). Exhaustive analysis of genome codon frequencies for microorganisms has revealed endogenous DNA of *E. coli* most commonly contains the CTG leucine-specifying codon, while the DNA of yeast and slime molds most commonly includes a TTA leucine-specifying codon. In view of this hierarchy, it is generally held that the likelihood of obtaining high levels of expression of a leucine-rich polypeptide by an *E. coli* host will depend to some extent on the frequency of codon use. For example, a gene rich in TTA codons will in all probability be poorly expressed in *E. coli*, whereas a CTG rich gene will probably highly express the polypeptide. Similarly, when yeast cells are the projected transformation host cells for expression of a leucine-rich polypeptide, a preferred codon for use in an inserted DNA would be TTA.

The implications of codon preference phenomena on recombinant DNA techniques are manifest, and the phenomenon may serve to explain many prior failures to achieve high expression levels of exogenous genes in successfully transformed host organisms--a less "preferred" codon may be repeatedly present in the inserted gene and the host cell machinery for expression may not operate as efficiently. This phenomenon suggests that synthetic genes which have been designed to include a projected host cell's preferred codons provide a preferred form of foreign

genetic material for practice of recombinant DNA techniques. Thus, one aspect of this invention is a vaccine administration protocol wherein the recombinant adenoviral vectors (prime and boost vectors) specifically include a gene which is codon optimized for expression in a human cellular environment. As noted herein, a preferred gene for use in the instant invention is a codon-optimized HIV gene and, particularly, HIV gag, pol, env, or nef although, as stated above, the adenoviral vehicles of the instant invention can utilize heterologous nucleic acid which may or may not be codon-optimized. In specific embodiments of the instant invention, the individual can be primed with an adenoviral vector comprising codon-optimized heterologous nucleic acid, and boosted with an adenovirus of an alternative serotype comprising non-codon-optimized nucleic acid. Administration of multiple antigens possesses the possibility for exploiting various different combinations of codon-optimized and non-codon-optimized sequences.

A vaccine composition comprising the recombinant viral vectors either in the priming or boosting dose in accordance with the instant invention may contain physiologically acceptable components, such as buffer, normal saline or phosphate buffered saline, sucrose, other salts and polysorbate. One preferred formulation has: 2.5-10 mM TRIS buffer, preferably about 5 mM TRIS buffer; 25-100 mM NaCl, preferably about 75 mM NaCl; 2.5-10% sucrose, preferably about 5% sucrose; 0.01 -2 mM $MgCl_2$; and 0.001%-0.01% polysorbate 80 (plant derived). The pH should range from about 7.0-9.0, preferably about 8.0. One skilled in the art will appreciate that other conventional vaccine excipients may also be used it make the formulation. The preferred formulation contains 5mM TRIS, 75 mM NaCl, 5% sucrose, 1mM $MgCl_2$, 0.005% polysorbate 80 at pH 8.0. This has a pH and divalent cation composition which is near the optimum for Ad5 and Ad6 stability and minimizes the potential for adsorption of virus to a glass surface. It does not cause tissue irritation upon intramuscular injection. It is preferably frozen until use.

The amount of viral particles in the vaccine composition to be introduced into a vaccine recipient will depend on the strength of the transcriptional and translational promoters used and on the immunogenicity of the expressed gene product. In general, an immunologically or prophylactically effective dose of $1x10^7$ to $1x10^{12}$ particles and preferably about $1x10^{10}$ to $1x10^{11}$ particles is administered directly into muscle tissue. Subcutaneous injection, intradermal introduction, impression through the skin, and other modes of administration such as intraperitoneal, intravenous, or inhalation

delivery are also contemplated. Parenteral administration, such as intravenous, intramuscular, subcutaneous or other means of administration of interleukin-12 protein, concurrently with or subsequent to parenteral introduction of the vaccine compositions of this invention is also advantageous.

The administration schemes of the instant invention are based on the priming of the immune response with an adenoviral vehicle of a first serotype comprising a gene encoding an HIV antigen (or antigens) and, following a predetermined length of time, boosting the adenovirus-primed response with an adenoviral vehicle of a second and alternative serotype comprising the gene encoding the HIV antigen(s). Multiple primings, typically, 1-4, are usually employed, although more may be used. The length of time between prime and boost may typically vary from about four months to a year, but other time frames may be used. The booster dose may be repeated at selected time intervals.

A large body of human and animal data supports the importance of cellular immune responses, especially CTL in controlling (or eliminating) HIV infection. In humans, very high levels of CTL develop following primary infection and correlate with the control of viremia. Several small groups of individuals have been described who are repeatedly exposed to HIV but remain uninfected; CTL has been noted in several of these cohorts. In the SIV model of HIV infection, CTL similarly develops following primary infection, and it has been demonstrated that addition of anti-CD8 monoclonal antibody abrogated this control of infection and leads to disease progression.

The following non-limiting Examples are presented to better illustrate the invention.

EXAMPLE 1

HIV-1 Gag Gene

A synthetic gene for HIV gag from HIV-1 strain CAM-1 was constructed using codons frequently used in humans; *see* Korber *et al.*, 1998 *Human Retroviruses and AIDS*, Los Alamos Nat'l Lab., Los Alamos, New Mexico; and Lathe, R., 1985 *J. Mol. Biol.* 183:1-12. Figure 2 illustrates the nucleotide sequence of the exemplified optimized codon version of full-length p55 gag. The gag gene of HIV-1 strain CAM-1 was selected as it closely resembles the consensus amino acid sequence for the clade

B (North American/European) sequence (Los Alamos HIV database). Advantage of this "codon-optimized" HIV gag gene as a vaccine component has been demonstrated in immunogenicity studies in mice. The "codon-optimized" HIV gag gene was shown to be over 50-fold more potent to induce cellular immunity than the wild type HIV gag gene when delivered as a DNA vaccine.

A KOZAK sequence (GCCACC) was introduced proceeding the initiating ATG of the gag gene for optimal expression. The HIV gag fragment with KOZAK sequence was amplified through PCR from V1Jns-HIV gag vector. PVIJnsHIVgag is a plasmid comprising the CMV immediate-early (IE) promoter and intron A, a full-length codon-optimized HIV gag gene, a bovine growth hormone-derived polyadenylation and transcriptional termination sequence, and a minimal pUC backbone; *see* Montgomery *et al.*, 1993 *DNA Cell Biol.* 12:777-783, for a description of the plasmid backbone.

EXAMPLE 2

Generation of Adenoviral Serotype 5 Vector Constructs

A. Removal of the Intron A Portion of the hCMV Promoter

GMP grade pVIJnsHIVgag was used as the starting material to amplify the hCMV promoter. The amplification was performed with primers suitably positioned to flank the hCMV promoter. A 5' primer was placed upstream of the *Msc*1 site of the hCMV promoter and a 3' primer (designed to contain the B*gl*II recognition sequence) was placed 3' of the hCMV promoter. The resulting PCR product (using high fidelity *Taq* polymerase) which encompassed the entire hCMV promoter (minus intron A) was cloned into TOPO PCR blunt vector and then removed by double digestion with *Msc*1 and B*gl*II. This fragment was then cloned back into the original GMP grade pV1JnsHIVgag plasmid from which the original promoter, intron A, and the gag gene were removed following *Msc*1 and *Bgl*II digestion. This ligation reaction resulted in the construction of a hCMV promoter (minus intron A) + bGHpA expression cassette within the original pV1JnsHIVgag vector backbone. This vector is designated pVIJnsCMV(no intron).

The FLgag gene was excised from pV1JnsHIVgag using *Bgl*II digestion and the 1,526 bp gene was gel purified and cloned into pV1JnsCMV(no intron) at the *Bgl*II site. Colonies were screened using *Sma*1 restriction enzymes to identify clones that carried the FLgag gene in the correct orientation. This plasmid, designated

pV1JnsCMV(no intron)-FLgag-bGHpA, was fully sequenced to confirm sequence integrity.

B. Construction of the Modified Shuttle Vector -"MRKpdelE1 Shuttle"

The modifications to the original Ad5 shuttle vector (pdelE1sp1A; a vector comprising Ad5 sequences from base pairs 1-341 and 3524-5798, with a multiple cloning region between nucleotides 341 and 3524 of Ad5, included the following three manipulations carried out in sequential cloning steps as follows:

(1) The left ITR region was extended to include the *Pac*1 site at the junction between the vector backbone and the adenovirus left ITR sequences. This allow for easier manipulations using the bacterial homologous recombination system.

(2) The packaging region was extended to include sequences of the wild-type (WT) adenovirus from 342 bp to 450 bp inclusive.

(3) The area downstream of pIX was extended 13 nucleotides (i.e., nucleotides 3511-3523 inclusive).

These modifications (Figure 4) effectively reduced the size of the E1 deletion without overlapping with any part of the E1A/E1B gene present in the transformed PER.C6® cell line. All manipulations were performed by modifying the Ad shuttle vector pdelE1sp1A.

Once the modifications were made to the shuttle vector, the changes were incorporated into the original Ad5 adenovector backbone pAdHVE3 by bacterial homologous recombination using *E. coli* BJ5183 chemically competent cells.

C. Construction of Modified Adenovector Backbone

An original adenovector pADHVE3 (comprising all Ad5 sequences except those nucleotides encompassing the E1 region) was reconstructed so that it would contain the modifications to the E1 region. This was accomplished by digesting the newly modified shuttle vector (MRKpdelE1 shuttle) with *Pac*1 and *Bst*Z1101 and isolating the 2,734 bp fragment which corresponds to the adenovirus sequence. This fragment was co-transformed with DNA from *Cla*1 linearized pAdHVE3 (E3+adenovector) into *E. coli* BJ5183 competent cells. At least two colonies from the transformation were selected and grown in Terrific™ broth for 6-8 hours until turbidity was reached. DNA was extracted from each cell pellet and then transformed into *E. coli* XL1 competent cells. One colony from the transformation was selected and grown for plasmid DNA purification. The plasmid was analyzed by restriction digestions to identify correct clones. The modified adenovector was designated

MRKpAdHVE3 (E3+ plasmid). Virus from the new adenovector (MRKHVE3) as well as the old version were generated in the PER.C6® cell lines. In addition, the multiple cloning site of the original shuttle vector contained ClaI , BamHI, Xho I, EcoRV, HindIII, Sal I, and Bgl II sites. This MCS was replaced with a new MCS containing Not I, Cla I, EcoRV and Asc I sites. This new MCS has been transferred to the MRKpAdHVE3 pre-plasmid along with the modification made to the packaging region and pIX gene.

D. Construction of the new shuttle vector containing modified gag transgene – "MRKpdelE1-CMV(no intron)-FLgag-bGHpA"

The modified plasmid pV1JnsCMV(no intron)-FLgag-bGHpA was digested with *Msc*1 overnight and then digested with *Sfi*1 for 2 hours at 50°C. The DNA was then treated with Mungbean nuclease for 30 minutes at 30°C. The DNA mixture was desalted using the Qiaex II kit and then Klenow treated for 30 minutes at 37°C to fully blunt the ends of the transgene fragment. The 2,559 bp transgene fragment was then gel purified. The modified shuttle vector (MRKpdelE1 shuttle) was linearized by digestion with EcoRV, treated with calf intestinal phosphatase and the resulting 6,479 bp fragment was then gel purified. The two purified fragments were then ligated together and several dozen clones were screened to check for insertion of the transgene within the shuttle vector. Diagnostic restriction digestion was performed to identify those clones carrying the transgene in the E1 parallel orientation.

E. Construction of the MRK FG Adenovector

The shuttle vector containing the HIV-1 gag transgene in the E1 parallel orientation, MRKpdelE1-CMV(no intron)-FLgag-bGHpA, was digested with *Pac*1. The reaction mixture was digested with *Bsf*Z171. The 5,291 bp fragment was purified by gel extraction. The MRKpAdHVE3 plasmid was digested with *Cla*1 overnight at 37°C and gel purified. About 100 ng of the 5,290 bp shuttle +transgene fragment and ~100 ng of linearized MRKpAdHVE3 DNA were co-transformed into *E. coli* BJ5183 chemically competent cells. Several clones were selected and grown in 2 ml Terrific™ broth for 6-8 hours, until turbidity was reached. The total DNA from the cell pellet was purified using Qiagen alkaline lysis and phenol chloroform method. The DNA was precipitated with isopropanol and resuspended in 20 µl dH_20. A 2 µl aliquot of this DNA was transformed into *E. coli* XL-1 competent cells. A single colony from the transformation was selected and grown overnight in 3 ml LB +100 µg/ml ampicillin. The DNA was isolated using Qiagen columns. A positive clone

was identified by digestion with the restriction enzyme *Bst*EII which cleaves within the gag gene as well as the plasmid backbone. The pre-plasmid clone is designated MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA and is 37,498 bp in size.

F. Virus generation of an enhanced adenoviral construct – "MRK Ad5 HIV-1gag"

MRK Ad5 HIV-1 gag contains the hCMV(no intron)-FLgag-bGHpA transgene inserted into the new E3+ adenovector backbone, MRKpAdHVE3, in the E1 parallel orientation. We have designated this adenovector MRK Ad5 HIV-1 gag. This construct was prepared as outlined below:

The pre-plasmid MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA was digested with *Pac*1 to release the vector backbone and 3.3 μg was transfected by the calcium phosphate method (Amersham Pharmacia Biotech.) in a 6 cm dish containing PER.C6® cells at ~60% confluence. Once CPE was reached (7-10 days), the culture was freeze/thawed three times and the cell debris pelleted. 1 ml of this cell lysate was used to infect into a 6 cm dish containing PER.C6® cells at 80-90% confluence. Once CPE was reached, the culture was freeze/thawed three times and the cell debris pelleted. The cell lysate was then used to infect a 15 cm dish containing PER.C6® cells at 80-90% confluence. This infection procedure was continued and expanded at passage 6. The virus was then extracted from the cell pellet by CsCl method. Two bandings were performed (3-gradient CsCl followed by a continuous CsCl gradient). Following the second banding, the virus was dialyzed in A105 buffer. Viral DNA was extracted using pronase treatment followed by phenol chloroform. The viral DNA was then digested with *Hind*III and radioactively labeled with [^{33}P]dATP. Following gel electrophoresis to separate the digestion products the gel was dried down on Whatman paper and then subjected to autoradiography. The digestion products were compared with the digestion products from the pre-plasmid (that had been digested with *Pac*1/*Hind*III prior to labeling). The expected sizes were observed, indicating that the virus had been successfully rescued.

EXAMPLE 3

Generation of Adenoviral Serotype 6 Vector Constructs

A. Construction of Ad6 Pre-Adenovirus Plasmid

An Ad6 based pre-adenovirus plasmid which could be used to generate first generation Ad6 vectors was constructed taking advantage of the extensive sequence

homology (approx. 98%) between Ad5 and Ad6. Homologous recombination was used to clone wtAd6 sequences into a bacterial plasmid.

The general strategy used to recover pAd6E1-E3+ as a bacterial plasmid is illustrated in Figure 7. Cotransformation of BJ 5183 bacteria with purified wt Ad6 viral DNA and a second DNA fragment termed the Ad5 ITR cassette resulted in the circularization of the viral genome by homologous recombination. The ITR cassette contains sequences from the right (bp 33798 to 35935) and left (bp 1 to 341 and bp 3525 to 5767) end of the Ad5 genome separated by plasmid sequences containing a bacterial origin of replication and an ampicillin resistance gene. The ITR cassette contains a deletion of E1 sequences from Ad5 342 to 3524. The Ad5 sequences in the ITR cassette provide regions of homology with the purified Ad6 viral DNA in which recombination can occur.

Potential clones were screened by restriction analysis and one clone was selected as pAd6E1-E3+. This clone was then sequenced in it entirety. pAd6E1-E3+ contains Ad5 sequences from bp 1 to 341 and from bp 3525 to 5548, Ad6 bp 5542 to 33784, and Ad5 bp 33967 to 35935 (bp numbers refer to the wt sequence for both Ad5 and Ad6). pAd6E1-E3+ contains the coding sequences for all Ad6 virion structural proteins which constitute its serotype specificity.

B. Construction of an Ad6 Pre-Adenovirus Plasmid containing the HIV-1 gag gene

(1) *Construction of Adenoviral Shuttle Vector:*

The shuttle plasmid MRKpdelE1(Pac/pIX/pack450)+CMVminFL-gag-BGHpA was constructed by inserting a synthetic full-length codon-optimized HIV-1 *gag* gene into MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.). MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) contains Ad5 sequences from bp 1 to 5792 with a deletion of E1 sequences from bp 451 to 3510. The HCMV promoter and BGH pA were inserted into the E1 deletion in an E1 parallel orientation with a unique BglII site separating them. The synthetic full-length codon-optimized HIV-1 *gag* gene was obtained from plasmid pV1Jns-HIV-FLgag-opt by *Bgl*II digestion, gel purified and ligated into the *Bgl*II restriction endonuclease site in MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.), generating plasmid MRKpdelE1(Pac/pIX/pack450)+CMVminFL-gag-BGHpA. The genetic structure of MRKpdelE1(Pac/pIX/pack450)+CMVminFL-gag-BGHpA was verified by PCR, restriction enzyme and DNA sequence analyses.

(2) *Construction of pre-adenovirus plasmid:*

Shuttle plasmid MRKpdelE1(Pac/pIX/pack450)+CMVminFL-gag-BGHpA was digested with restriction enzymes *Pac*1 and *Bst*1107I and then co-transformed into *E. coli* strain BJ5183 with linearized (*Cla*I-digested) adenoviral backbone plasmid, pAd6E1-E3+. The genetic structure of the resulting pMRKAd6gag was verified by restriction enzyme and DNA sequence analysis. The vectors were transformed into competent *E. coli* XL-1 Blue for large-scale production. The recovered plasmid was verified by restriction enzyme digestion and DNA sequence analysis, and by expression of the gag transgene in transient transfection cell culture.

pMRKAd6gag contains Ad5 bp 1 to 450 and from bp 3511 to 5548, Ad6 bp 5542 to 33784, and Ad5 bp 33967 to 35935 (bp numbers refer to the wt sequence for both Ad5 and Ad6). In the plasmid the viral ITRs are joined by plasmid sequences that contain the bacterial origin of replication and an ampicillin resistance gene.

C. Generation of research-grade recombinant MRKAd6gag

To prepare virus for pre-clinical immunogenicity studies, the pre-adenovirus plasmid pMRKAd6gag was rescued as infectious virions in PER.C6® adherent monolayer cell culture. To rescue infectious virus, 10 μg of pMRKAd6gag was digested with restriction enzyme *Pac*I (New England Biolabs) and transfected into a 6 cm dish of PER.C6® cells using the calcium phosphate co-precipitation technique (Cell Phect Transfection Kit, Amersham Pharmacia Biotech Inc.). *Pac*I digestion releases the viral genome from plasmid sequences allowing viral replication to occur after entry into PER.C6®cells. Infected cells and media were harvested after complete viral cytopathic effect (CPE) was observed. The virus stock was amplified by multiple passages in PER.C6® cells. At the final passage virus was purified from the cell pellet by CsCl ultracentrifugation. The identity and purity of the purified virus was confirmed by restriction endonuclease analysis of purified viral DNA and by gag ELISA of culture supernatants from virus infected mammalian cells grown in vitro. For restriction analysis, digested viral DNA was end-labeled with P^{33}-dATP, size-fractionated by agarose gel electrophoresis, and visualized by autoradiography.

All viral constructs (adenovirus serotypes 5 and 6) were confirmed for Gag expression by Western blot analysis.

EXAMPLE 4

Immunization

Rhesus macaques were between 3-10 kg in weight. In all cases, the total dose of each vaccine was suspended in 1 mL of buffer. The macaques were anesthetized (ketamine/xylazine) and the vaccines were delivered intramuscularly ("i.m.") in 0.5-mL aliquots into both deltoid muscles using tuberculin syringes (Becton-Dickinson, Franklin Lakes, NJ). Peripheral blood mononuclear cells (PBMC) were prepared from blood samples collected at several time points during the immunization regimen. All animal care and treatment were in accordance with standards approved by the Institutional Animal Care and Use Committee according to the principles set forth in the *Guide for Care and Use of Laboratory Animals*, Institute of Laboratory Animal Resources, National Research Council.

EXAMPLE 5

ELISPOT Assay

The IFN-γ ELISPOT assays for rhesus macaques were conducted following a previously described protocol (Allen *et al.*, 2001 *J. Virol.* 75(2):738-749), with some modifications. For antigen-specific stimulation, a peptide pool was prepared from 20-amino acid ("aa") peptides that encompass the entire HIV-1 gag sequence with 10-aa overlaps (Synpep Corp., Dublin, CA). To each well, 50 μL of 2-4 x 10^5 peripheral blood mononuclear cells (PBMCs) were added. The cells were counted using Beckman Coulter Z2 particle analyzer with a lower size cut-off set at 80 femtoliters ("fL"). Either 50 μL of media or the gag peptide pool at 8 μg/mL concentration per peptide were added to the PBMC. The samples were incubated at 37°C, 5% CO_2 for 20-24 hrs. Spots were developed accordingly and the plates were processed using custom-built imager and automatic counting subroutine based on the ImagePro platform (Silver Spring, MD). The counts were normalized to 10^6 cell input.

EXAMPLE 6

Anti-p24 ELISA

A modified competitive anti-p24 assay was developed using reagents from the Coulter p24 Antigen Assay kit (Beckman Coulter, Fullerton, CA). Briefly, to a 250-μL serum sample, 20 μL of Lyse Buffer and 15 μL of p24 antigen (9.375 pg) from the Coulter kit were added. After mixing, 200 μL of each sample were added to wells

coated with a mouse anti-p24 mAb from the Coulter kit and incubated for 1.5 hr at 37°C. The wells were then washed and 200 μL of Biotin Reagent (polyclonal anti-p24-biotin) from the Coulter kit was added to each well. After a 1 hr, 37°C incubation, detection was achieved using strepavidin-conjugated horseradish peroxidase and TMB substrate as described in the Coulter Kit. OD_{450nm} values were recorded. A 7-point standard curve was generated using a serial 2-fold dilution of serum from an HIV-seropositive individual. The lower cut-off for the assay is arbitrarily set at 10 milli Merck units/mL (mMU/mL) defined by a dilution of the seropositive human serum. This cutoff falls at approximately 65% of the maximum bound control signal which corresponds to that obtained with the diluent control only and with no positive analyte.

EXAMPLE 7

Intracellular Cytokine Staining

To 1 ml of 2 x 10^6 PBMC/mL in complete RPMI media (in 17x100mm round bottom polypropylene tubes (Sarstedt, Newton, NC)), anti-hCD28 (clone L293, Becton-Dickinson) and anti-hCD49d (clone L25, Becton-Dickinson) monoclonal antibodies were added to a final concentration of 1 μg/mL. For gag-specific stimulation, 10 μL of the peptide pool (at 0.4 mg/mL per peptide) were added. The tubes were incubated at 37 °C for 1 hr., after which 20 μL of 5 mg/mL of brefeldin A (Sigma) were added. The cells were incubated for 16 hours at 37 °C, 5% CO_2, 90% humidity. 4 mL cold PBS/2%FBS were added to each tube and the cells were pelleted for 10 min at 1200 rpm. The cells were re-suspended in PBS/2%FBS and stained (30 min, 4 °C) for surface markers using several fluorescent-tagged mAbs: 20 μL per tube anti-hCD3-APC, clone FN-18 (Biosource); 20 μL anti-hCD8-PerCP, clone SK1 (Becton Dickinson); and 20 μL anti-hCD4-PE, clone SK3 (Becton Dickinson). Sample handling from this stage was conducted in the dark. The cells were washed and incubated in 750 μL 1xFACS Perm buffer (Becton Dickinson) for 10 minutes at room temperature. The cells were pelleted and re-suspended in PBS/2%FBS and 0.1 μg of FITC-anti-hIFN-γ, clone MD-1 (Biosource) was added. After 30 minutes of incubation, the cells were washed and re-suspended in PBS. Samples were analyzed using all four color channels of the Becton Dickinson FACS Calibur instrument. To analyze the data, the low side- and forward-scatter lymphocyte population was initially gated and a common fluorescence cut-off for

cytokine-positive events was used for both $CD4^+$ and $CD8^+$ populations, and for both mock and gag-peptide reaction tubes of a sample.

EXAMPLE 8

Results

A. Immunization Regimen

Cohorts of 3-6 rhesus macaques were immunized following homologous and heterologous prime-boost regimens involving MRKAd5 and MRKAd6 vectors expressing the same codon-optimized HIV-1 gag. The immunization schedule is described in Table 1.

Table 1.

Group	Prime	Boost (month 6)
1	10e9 vp MRKAd5-HIVgag at week 0, 4	10e9 vp MRKAd5-HIVgag
2	10e9 vp MRKAd6-HIVgag at week 0, 4	10e9 vp MRKAd6-HIVgag
3	10e9 vp MRKAd5-HIVgag at week 0, 4	10e9 pfu MRKAd6-HIVgag

B. T Cell Immune Responses

Vaccine-induced T cell responses against HIV-1 gag were quantified using IFN-gamma ELISPOT assay against a pool of 20-aa peptides that encompassed the entire protein sequence. The results are shown in Figure 5. They are expressed as the number of spot-forming cells (SFC) per million peripheral blood mononuclear cells (PBMCs) that responded to the peptide pool minus the mock control.

The Figure shows the T cell responses induced by two priming immunizations with 10e9 vp MRKAd5-HIVgag followed by a 10e9 vp MRKAd5-HIVgag booster after a long rest (a period of 20-23 weeks; 22 for the MRKAd6-MRKAd6 subjects; 22 for subjects 99D262, 99C117, and 99D227 of the MRKAd5-MRKAd5 group; and 23 for the remaining subjects). Administration of the same dose of MRKAd5 HIV-1 gag at approximately month 6 resulted in slight increases compared to the levels just prior to the boost; the post-boost levels were largely comparable to if not weaker than the peak levels before the boost. This is possibly due to the presence of neutralizing immunity generated against the vector by the first two immunizations. The responses after the boost did not surpass 500 gag-specific T cells per 10e6 PBMC, with a mean of 275 SFC/10e6 PBMC for all 6 monkeys. Similar results were observed when monkeys were given three of 10e9 vp MRKAd6 HIV-1 gag (at 0, 1, 6 months). In two out of the three monkeys, the post-boost levels did not surpass 500 SFC/10e6

PBMC. In contrast, when both modalities are combined in which animals were given two priming doses of 10e9 vp MRKAd5-HIVgag and a single booster shot of 10e9 pfu MRKAd6-HIVgag, the levels of gag-specific T cells increased to peak responses above 1000 SFC/10e6 PBMC for all 3 monkeys. The ability of MRKAd6-HIVgag to boost effectively MRKAd5-gag-primed immune responses more effectively is possibly due to the presence of neutralizing immunity generated against the MRKAd5 vector by the first two immunizations. The ability of Ad6 to boost primed responses was also evident using a lower priming dose of 10^7 vp of MRKAd5 HIV-1 gag (Figure 6).

PBMCs from the vaccinees of the heterologous MRKAd5 prime-MRKAd6 boost regimen were analyzed for intracellular IFN-γ staining after the priming immunizations (wk 13) and after the booster immunizations (wk 31). The assay provided information on the relative amounts of $CD4^+$ and $CD8^+$ gag-specific T cells in the peripheral blood (Table 2). The results indicated that heterologous prime-boost immunization approach was able to elicit in rhesus macaques both HIV-specific CD4+ and CD8+ T cells.

Table 2.

			Post Prime		Post Boost	
Prime	Boost	ID	%CD4+	%CD8+	%CD4+	%CD8+
MRKAd5-HIVgag	MRKAd6-HIVgag	99C216	0.05	0.21	0.10	1.45
10^9 vp	10^9 pfu	99C231	0.03	0.10	0.16	1.41
wk 0, 4	wk 27	99C132	0.00	0.02	0.04	0.15

Numbers reflect the percentages of circulating CD3+ lymphocytes that are either gag-specific CD4+ or gag-specific CD8+ cells. Mocks values have been subtracted.
*No detectable antigen-specific CD4+ T cells above background
**Collected at wk 35 instead of wk 31

C. Humoral Immune Responses

The p24-specific antibody titers were determined for each animal at several time points. The geometric mean titers for each cohort were calculated and shown in Figure 10. Two doses of MRKAd5 HIV-1 gag or MRKAd6 HIV-1 gag were able to induce moderate levels of anti-p24 antibodies (about 1000 mMU/mL). Administration of the same viral vector booster resulted in 5-10 fold increase in the humoral immune responses. Boosting MRKAd5 HIV-1 gag primed monkeys with MRKAd6-gag resulted in a comparable in antibody levels. Boosting with the same virus can have its limitations, though, as the effect can be negatively impacted by any

significant neutralizing Ad5-specific activity. The booster effect of a non-matched Ad serotype, by contrast, would not be affected by any pre-existing neutralizing titers directed at Ad5.

EXAMPLE 9

Generation of a Completely Adenoviral Serotype 6 Vector Construct

A. Construction of a Completely Ad6 Pre-Adenovirus Plasmid

An Ad6 based pre-adenovirus plasmid derived from Ad6 sequence and not constructed taking advantage of the homology between Ad5 and Ad6 can be generated and used to generate first generation Ad6 vectors. Homologous recombination is used to clone wtAd6 sequences into a bacterial plasmid.

The general strategy used to recover such a pMRKAd6E1- bacterial plasmid is illustrated in Figure 13. Basically, cotransformation of BJ 5183 bacteria with purified wt Ad6 viral DNA and a second DNA fragment termed the Ad6 ITR cassette would effectuate circularization of the viral genome by homologous recombination. The ITR cassette contains sequences from the right (bp 35460 to 35759) and left (bp 1 to 450 and bp 3508 to 3807) end of the Ad6 genome separated by plasmid sequences containing a bacterial origin of replication and an ampicillin resistance gene. These three segments were generated by PCR and cloned sequentially into pNEB193 (a commonly used commercially available cloning plasmid (New England Biolabs cat# N3051S) containing a bacterial origin of replication ,ampicillin resistance gene and a multiple cloning site into which the PCR products are introduced), generating pNEBAd6-3 (the ITR cassette). The ITR cassette contains a deletion of E1 sequences from Ad5 451 to 3507. The Ad6 sequences in the ITR cassette provide regions of homology with the purified Ad6 viral DNA in which recombination can occur.

PMRKAd6E1- can then be used to generate first generation Ad6 vectors containing transgenes in E1 as described in the previous example.

EXAMPLE 10

In Vivo Immunogenicity

A. Immunization

Rhesus macaques were between 3-10 kg in weight. In all cases, the total dose of each vaccine was suspended in 1 mL of buffer. The macaques were anesthetized

(ketamine/xylazine) and the vaccines were delivered i.m. in 0.5-mL aliquots into both deltoid muscles using tuberculin syringes (Becton-Dickinson, Franklin Lakes, NJ). Peripheral blood mononuclear cells (PBMC) were prepared from blood samples collected at several time points during the immunization regimen. All animal care and treatment were in accordance with standards approved by the Institutional Animal Care and Use Committee according to the principles set forth in the *Guide for Care and Use of Laboratory Animals*, Institute of Laboratory Animal Resources, National Research Council.

B. ELISPOT Assay

The IFN-γ ELISPOT assays for rhesus macaques were conducted following a previously described protocol (Allen et al., 2001 *J. Virol.* 75(2): 738-749), with some modifications. For antigen-specific stimulation, a peptide pool was prepared from 20-aa peptides that encompass the entire HIV-1 gag sequence with 10-aa overlaps (Synpep Corp., Dublin, CA). To each well, 50 μL of 2-4 x 10^5 peripheral blood mononuclear cells (PBMCs) were added; the cells were counted using Beckman Coulter Z2 particle analyzer with a lower size cut-off set at 80 fL. Either 50 μL of media or the gag peptide pool at 8 μg/mL concentration per peptide were added to the PBMC. The samples were incubated at 37°C, 5% CO_2 for 20-24 hrs. Spots were developed accordingly and the plates were processed using custom-built imager and automatic counting subroutine based on the ImagePro platform (Silver Spring, MD); the counts were normalized to 10^6 cell input.

C. Results

Rare Serotype Vaccine Vector as a Heterologous Booster. A cohort of three rhesus macaques was immunized initially with 3 doses (wk 0, 4, 16) of 10^8 vp of MRKAd5-gag. At wk 59, the animals received a booster vaccine of 10^{10} vp Ad35ΔE1gagΔE4Ad5Orf6 (an Ad35 virus engineered to contain an E1 deletion (from Ad35 bps 457-3402); and a deletion of E4 Orf6 (from Ad35 bps 31912-34418) substituted with Ad5 Orf6). A separate cohort of naïve animals received a single dose of the booster vaccine. The results of the IFN-γ ELISPOT analyses of PBMC collected during the course of the studies are shown in Table 3.

Table 3.

Animal	Prime (Wk 0, 4, 16)	Boost (Wk 59)	Pre		Prime[b]		Pre-Boost[c]		Post-Boost[d]	
			Mock[a]	Gag[a]	Mock	Gag	Mock	Gag	Mock	Gag
Monkey 11	10^8 vp MRKAd5-gag	10^{10} vp Ad35ΔE1gagΔE4Ad5Orf6	0	1	1	153	0	25	3	1120
Monkey 12	10^8 vp MRKAd5-gag	10^{10} vp Ad35ΔE1gagΔE4Ad5Orf6	4	6	3	269	0	23	1	659
Monkey 13	10^8 vp MRKAd5-gag	10^{10} vp Ad35ΔE1gagΔE4Ad5Orf6	1	3	3	150	0	10	1	489
Monkey 14	none	10^{10} vp Ad35ΔE1gagΔE4Ad5Orf6	1	9	ND[e]	ND	ND	ND	0	20
Monkey 15	none	10^{10} vp Ad35ΔE1gagΔE4Ad5Orf6	3	3	ND	ND	ND	ND	1	81
Monkey 16	none	10^{10} vp Ad35ΔE1gagΔE4Ad5Orf6	0	6	ND	ND	ND	ND	0	46

[a]Mock, no peptide: gag, 20-mer peptide pool encompassing entire gag sequence
[b]Peak response after 2 or 3 doses of the priming vaccine
[c]Wk 59
[d]4 wks after boost
[e]ND, not determined

It is apparent that Ad35-based HIV vectors can be utilized to amplify the existing pools of HIV-specific T cells. The increases in the levels of gag-specific T cells from the pre-boost levels to those measured at 4 wks post boost were consistently larger than the levels induced by the same booster vaccine in naïve animals.

WHAT IS CLAIMED IS:

1. A method for inducing an enhanced immunological response against an HIV-1 antigen in a mammalian host, said method comprising the steps of:

(a) inoculating the mammalian host with a recombinant adenoviral vector of a first serotype which is at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding an HIV-1 antigen or immunologically relevant modification thereof; and thereafter

(b) inoculating the mammalian host with a boosting immunization comprising a recombinant adenoviral vector of a second serotype which is at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding the HIV-1 antigen or immunologically relevant modification thereof.

2. A method in accordance with claim 1 wherein the HIV-1 antigen is HIV-1 gag.

3. A method in accordance with claim 1 wherein the HIV-1 antigen is HIV-1 nef.

4. A method in acccordance with claim 1 wherein the HIV-1 antigen is HIV-1 pol.

5. A method in accordance with claim 1 wherein at least one gene encoding the HIV-1 antigen or immunologically relevant modification thereof comprises codons optimized for expression in a mammalian host.

6. A method in accordance with claim 1 wherein one or more of the recombinant adenoviral vectors comprise a gene expression cassette, said gene expression cassette which comprises:

(a) a nucleic acid encoding an HIV-1 antigen;

(b) a heterologous promoter operatively linked to the nucleic acid encoding the antigen; and

(c) a transcription termination sequence.

7. A method in accordance with claim 6 wherein the gene expression cassette in at least one of the recombinant adenoviral vectors is inserted into the E1 region.

8. A method in accordance with claim 6 wherein the promoter is an immediate early human cytomegalovirus promoter.

9. A method in accordance with claim 6 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.

10. A method for inducing an enhanced immunological response against an HIV-1 antigen in a mammalian host, said method comprising the steps of:

(a) inoculating the mammalian host with a recombinant adenoviral vector of serotype 5 at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding an HIV-1 antigen or immunologically relevant modification thereof; and thereafter

(b) inoculating the mammalian host with a boosting immunization comprising a recombinant adenoviral vector of serotype 6 at least partially deleted in E1 and

devoid of E1 activity comprising a gene encoding the HIV-1 antigen or immunologically relevant modification thereof.

11. A method in accordance with claim 10 wherein the recombinant adenoviral vector of step (a) is deleted of base pairs 451-3510.

12. A method in accordance with claim 10 wherein the recombinant adenoviral vector of step (b) is deleted of base pairs 451-3507.

13. A method in accordance with claim 10 wherein at least one gene encoding the HIV-1 antigen or immunologically relevant modification thereof comprises codons optimized for expression in a mammalian host.

14. A method in accordance with claim 10 wherein the HIV-1 antigen is HIV-1 gag.

15. A method in accordance with claim 10 wherein the HIV-1 antigen is HIV-1 nef.

16. A method in acccordance with claim 10 wherein the HIV-1 antigen is HIV-1 pol.

17. A method in accordance with claim 10 wherein one or more of the recombinant adenoviral vectors comprise a gene expression cassette, said gene expression cassette which comprises:

(a) a nucleic acid encoding an HIV-1 antigen;

(b) a heterologous promoter operatively linked to the nucleic acid encoding the antigen; and

(c) a transcription termination sequence.

18. A method in accordance with claim 17 wherein the gene expression cassette in at least one of the recombinant adenoviral vectors is inserted into the E1 region.

19. A method in accordance with claim 17 wherein the promoter is an immediate early human cytomegalovirus promoter.

20. A method in accordance with claim 17 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.

21. A method for inducing an enhanced immunological response against an HIV-1 gag antigen in a mammalian host, said method comprising the steps of:

(a) inoculating the mammalian host with a recombinant adenoviral vector of serotype 5 at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding an HIV-1 gag antigen or immunologically relevant modification thereof; and thereafter

(b) inoculating the mammalian host with a boosting immunization comprising a recombinant adenoviral vector of serotype 6 at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding the HIV-1 gag antigen or immunologically relevant modification thereof.

22. A method for inducing an enhanced immunological response against an HIV-1 antigen in a mammalian host, said method comprising the steps of:

(a) inoculating the mammalian host with a recombinant adenoviral vector of serotype 5 at least partially deleted in E1 and devoid of E1 activity comprising a gene

encoding an HIV-1 antigen or immunologically relevant modification thereof; and thereafter

(b) inoculating the mammalian host with a boosting immunization comprising a recombinant adenoviral vector of serotype 35 at least partially deleted in E1 and devoid of E1 activity comprising a gene encoding the HIV-1 antigen or immunologically relevant modification thereof.

23. A method in accordance with claim 22 wherein at least one gene encoding the HIV-1 antigen or immunologically relevant modification thereof comprises codons optimized for expression in a mammalian host.

24. A method in accordance with claim 22 wherein the HIV-1 antigen is HIV-1 gag.

25. A method in accordance with claim 22 wherein the HIV-1 antigen is HIV-1 nef.

26. A method in acccordance with claim 22 wherein the HIV-1 antigen is HIV-1 pol.

27. A method in accordance with claim 22 wherein one or more of the recombinant adenoviral vectors comprise a gene expression cassette, said gene expression cassette which comprises:

(a) a nucleic acid encoding an HIV-1 antigen;

(b) a heterologous promoter operatively linked to the nucleic acid encoding the antigen; and

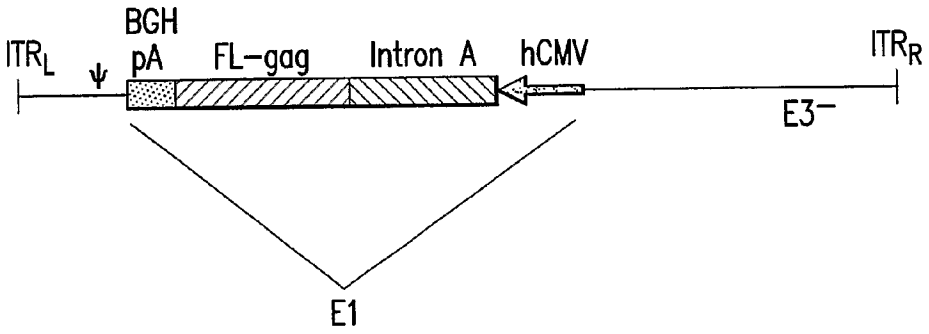
(c) a transcription termination sequence.

28. A method in accordance with claim 27 wherein the gene expression cassette in at least one of the recombinant adenoviral vectors is inserted into the E1 region.

29. A method in accordance with claim 27 wherein the promoter is an immediate early human cytomegalovirus promoter.

30. A method in accordance with claim 27 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.

ORIGINAL ADENOVECTOR CONSTRUCT:



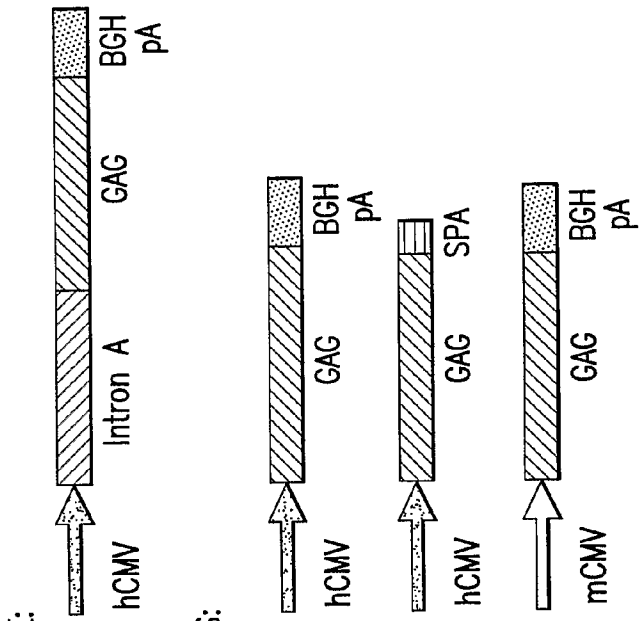
ORIGINAL HIV-1 gag ADENOVECTOR.

FIG.1

Sequence of the open reading frame for FL-gag (human codon optimized)

atgggtgctagggcttctgtgctgtctggtggtgagctggacaagtgggagaagatcaggctgaggcctggtgg
caagaagaagtacaagctaaagcacattgtgtgggcctccagggagctggagaggtttgctgtgaaccctggc
ctgctggagacctctgaggggtgcaggcagatcctgggccagctccagccctccctgcaaacaggctctgagg
agctgaggtccctgtacaacacagtggctaccctgtactgtgtgcaccagaagattgatgtgaaggacaccaag
gaggccctggagaagattgaggaggagcagaacaagtccaagaagaaggcccagcaggctgctgctggc
acaggcaactccagccaggtgtcccagaactaccccattgtgcagaacctccagggccagatggtgcaccag
gccatctcccccccggaccctgaatgcctgggtgaaggtggtggaggagaaggccttctcccctgaggtgatccc
catgttctctgccctgtctgagggtgccacccccccaggacctgaacaccatgctgaacacagtgggggccatc
aggctgccatgcagatgctgaaggagaccatcaatgaggaggctgctgagtgggacaggctgcatcctgtgc
acgctggccccattgcccccggccagatgagggagcccaggggctctgacattgctggcaccacctccaccct
ccaggagcagattggctggatgaccaacaacccccccatccctgtgggggaaatctacaagaggtggatcat
cctgggcctgaacaagattgtgaggatgtactcccccacctccatcctggacatcaggcagggccccaaggag
cccttcagggactatgtggacaggttctacaagaccctgagggctgagcaggcctcccaggaggtgaagaact
ggatgacagagaccctgctggtgcagaatgccaaccctgactgcaagaccatcctgaaggccctgggccctg
ctgccaccctggaggagatgatgacagcctgccaggggtgggggccctggtcacaaggccagggtgctg
gctgaggccatgtcccaggtgaccaactccgccaccatcatgatgcagaggggcaacttcaggaaccagag
gaagacagtgaagtgcttcaactgtggcaaggtgggccacattgccaagaactgtagggcccccaggaaga
agggctgctggaagtgtggcaaggagggccaccagatgaaggactgcaatgagaggcaggccaacttcctg
ggcaaaatctggccctcccacaagggcaggcctggcaacttcctccagtccaggcctgagcccacagcccct
cccgaggagtccttcaggtttggggaggagaagaccacccccagccagaagcaggagcccattgacaagg
agctgtacccccctggcctccctgaggtccctgtttggcaacgacccctcctcccagtaaaataaagcccgggca
gat

FIG.2



DIAGRAMMATIC REPRESENTATION OF THE ORIGINAL HIV-1 GAG TRANSGENE AND THE SERIES OF NEW TRANSGENE CONSTRUCTIONS.

FIG.3

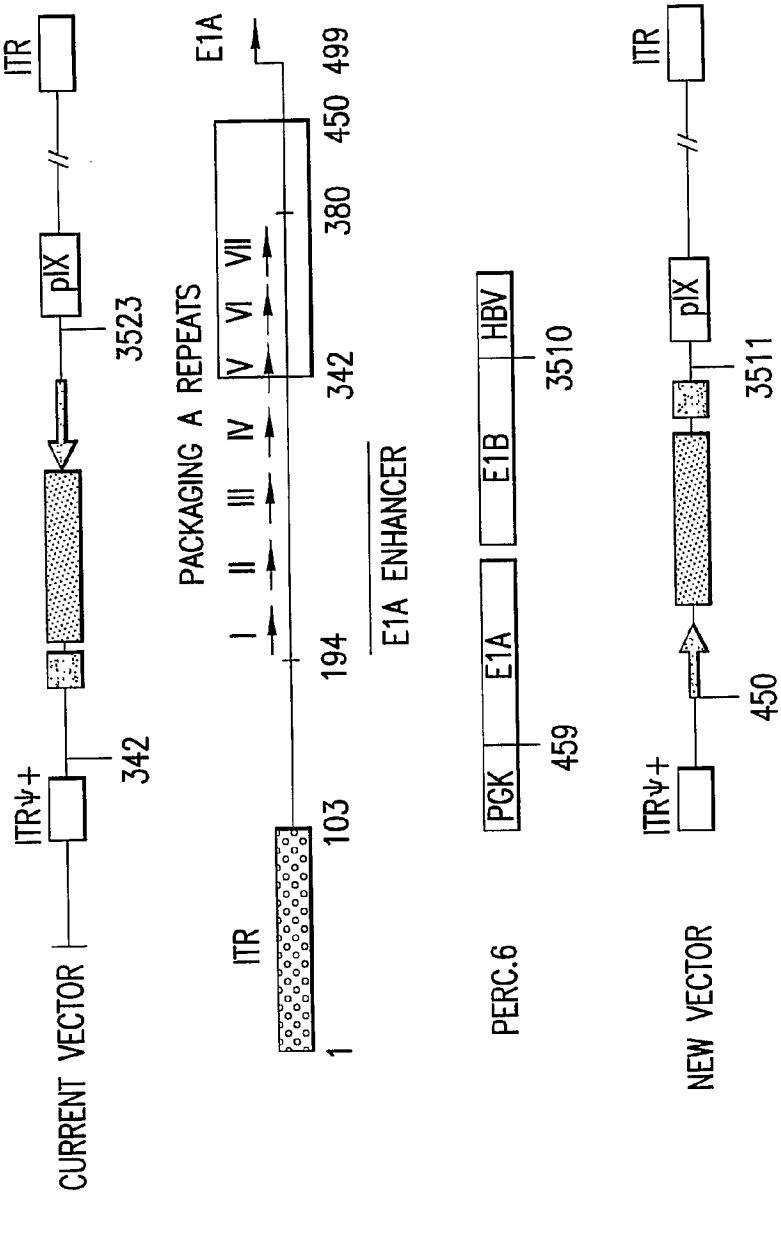


FIG.4

MODIFICATIONS MADE TO THE CURRENT ADENOVECTOR BACKBONE IN THE GENERATION OF THE NEW VECTOR.

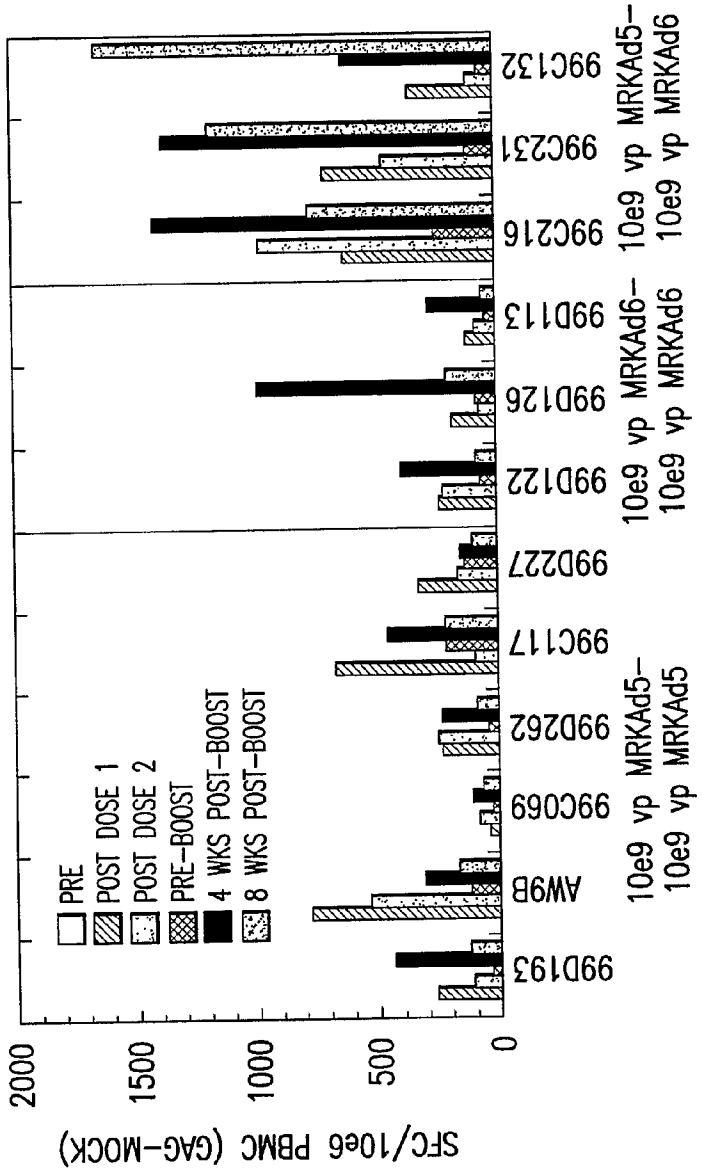

SFC/10e6 PBMC (GAG-MOCK)
2000
1500
1000
500
0
PRE
POST DOSE 1
POST DOSE 2
PRE-BOOST
4 WKS POST-BOOST
8 WKS POST-BOOST
99D193
AW9B
99C069
99D262
99C117
99D227
99D122
99D126
99D113
99C216
99C231
99C132
10e9 vp MRKAd5-
10e9 vp MRKAd5
10e9 vp MRKAd6-
10e9 vp MRKAd6
10e9 vp MRKAd5-
10e9 vp MRKAd6

FIG.5

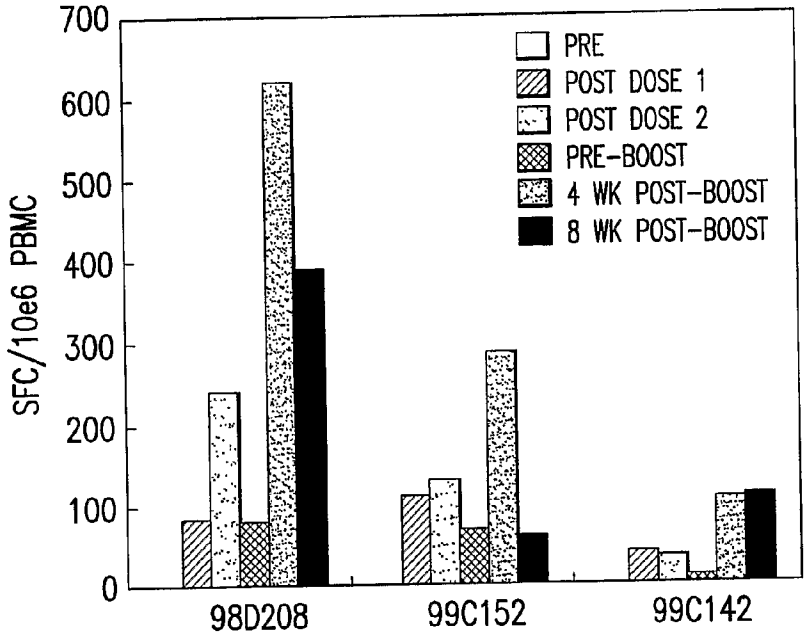

700
600
500
400
300
200
100
0
SFC/10e6 PBMC
PRE
POST DOSE 1
POST DOSE 2
PRE-BOOST
4 WK POST-BOOST
8 WK POST-BOOST
98D208
99C152
99C142

FIG.6

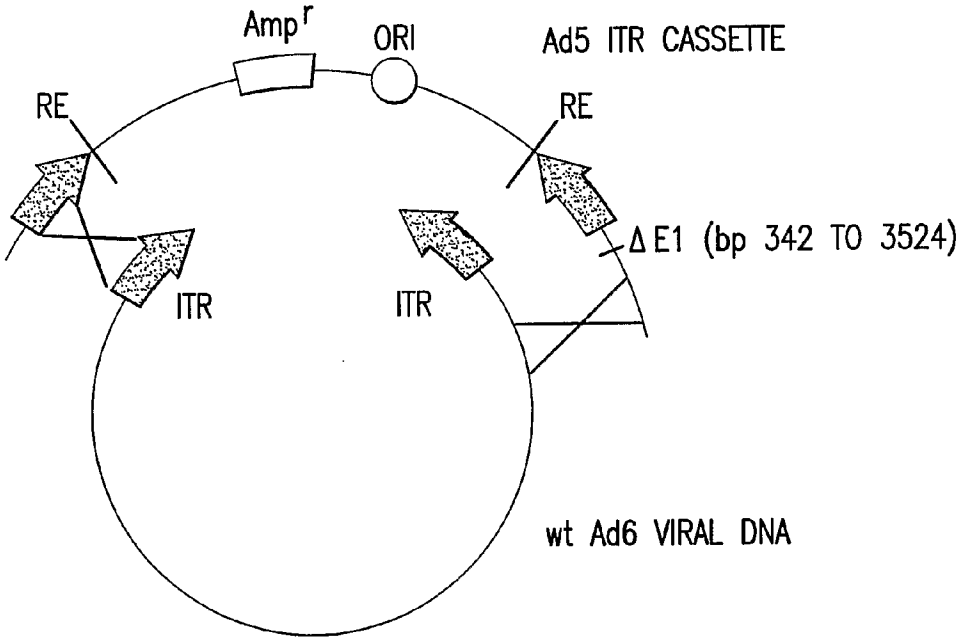

Amp r
ORI
Ad5 ITR CASSETTE
RE
RE
Δ E1 (bp 342 TO 3524)
ITR
ITR
wt Ad6 VIRAL DNA

FIG.7

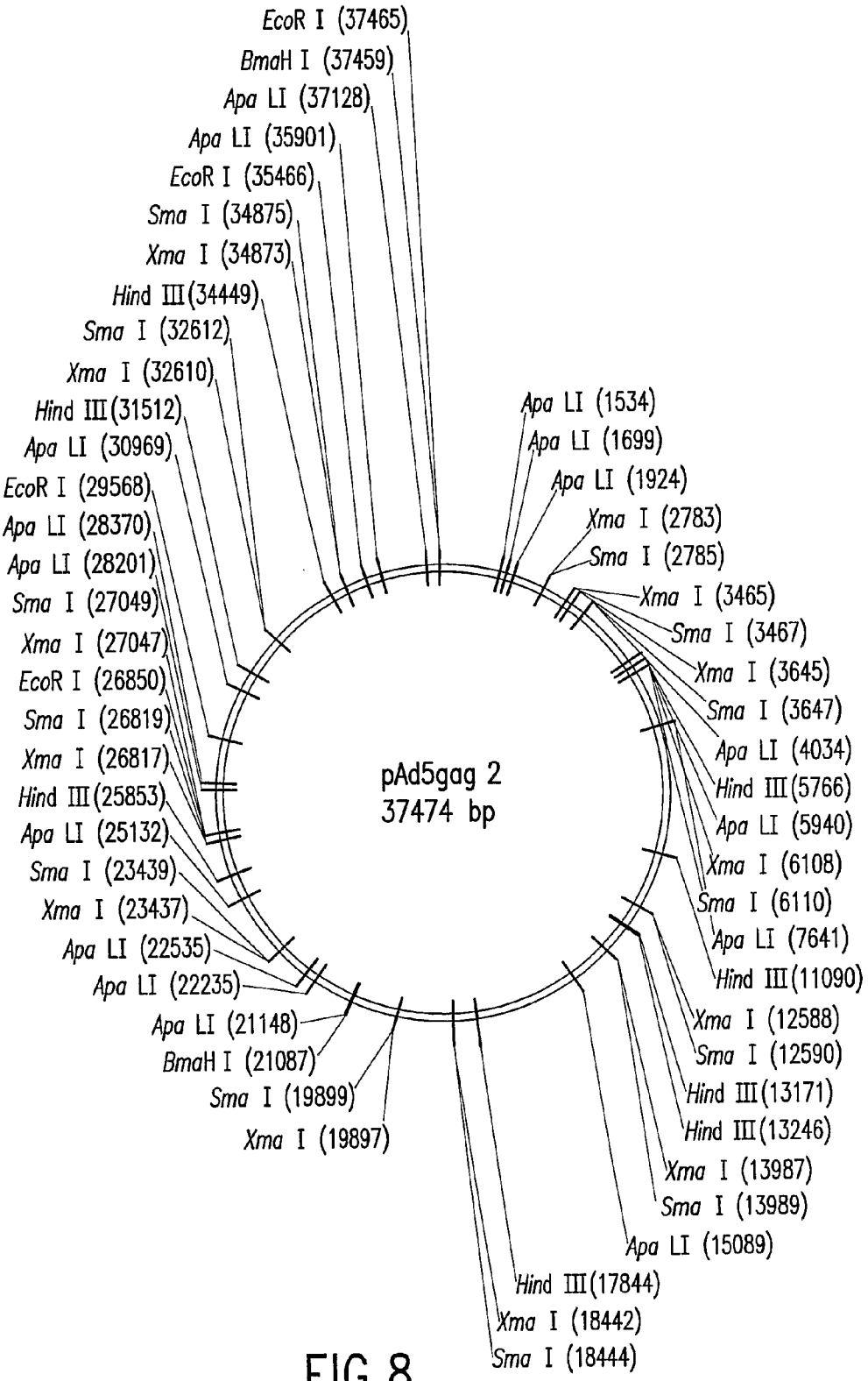

EcoR I (37465)
BmaH I (37459)
Apa LI (37128)
Apa LI (35901)
EcoR I (35466)
Sma I (34875)
Xma I (34873)
Hind III(34449)
Sma I (32612)
Xma I (32610)
Hind III(31512)
Apa LI (30969)
EcoR I (29568)
Apa LI (28370)
Apa LI (28201)
Sma I (27049)
Xma I (27047)
EcoR I (26850)
Sma I (26819)
Xma I (26817)
Hind III(25853)
Apa LI (25132)
Sma I (23439)
Xma I (23437)
Apa LI (22535)
Apa LI (22235)
Apa LI (21148)
BmaH I (21087)
Sma I (19899)
Xma I (19897)
Apa LI (1534)
Apa LI (1699)
Apa LI (1924)
Xma I (2783)
Sma I (2785)
Xma I (3465)
Sma I (3467)
Xma I (3645)
Sma I (3647)
Apa LI (4034)
Hind III(5766)
Apa LI (5940)
Xma I (6108)
Sma I (6110)
Apa LI (7641)
Hind III(11090)
Xma I (12588)
Sma I (12590)
Hind III(13171)
Hind III(13246)
Xma I (13987)
Sma I (13989)
Apa LI (15089)
Hind III(17844)
Xma I (18442)
Sma I (18444)
pAd5gag 2
37474 bp

FIG.8

PacI

```
   1 TTCTTAATTA ACATCATCAA TAATATACCT TATTTTGGAT TGAAGCCAAT
     AAGAATTAAT TGTAGTAGTT ATTATATGGA ATAAAACCTA ACTTCGGTTA

  51 ATGATAATGA GGGGGTGGAG TTTGTGACGT GGCGCGGGGC GTGGGAACGG
     TACTATTACT CCCCCACCTC AAACACTGCA CCGCGCCCCG CACCCTTGCC

 101 GGCGGGTGAC GTAGTAGTGT GGCGGAAGTG TGATGTTGCA AGTGTGGCGG
     CCGCCCACTG CATCATCACA CCGCCTTCAC ACTACAACGT TCACACCGCC

 151 AACACATGTA AGCGACGGAT GTGGCAAAAG TGACGTTTTT GGTGTGCGCC
     TTGTGTACAT TCGCTGCCTA CACCGTTTTC ACTGCAAAAA CCACACGCGG

 201 GGTGTACACA GGAAGTGACA ATTTTCGCGC GGTTTTAGGC GGATGTTGTA
     CCACATGTGT CCTTCACTGT TAAAAGCGCG CCAAAATCCG CCTACAACAT

 251 GTAAATTTGG GCGTAACCGA GTAAGATTTG GCCATTTTCG CGGGAAAACT
     CATTTAAACC CGCATTGGCT CATTCTAAAC CGGTAAAAGC GCCCTTTTGA

 301 GAATAAGAGG AAGTGAAATC TGAATAATTT TGTGTTACTC ATAGCGCGTA
     CTTATTCTCC TTCACTTTAG ACTTATTAAA ACACAATGAG TATCGCGCAT

 351 ATATTTGTCT AGGGCCGCGG GGACTTTGAC CGTTTACGTG GAGACTCGCC
     TATAAACAGA TCCCGGCGCC CCTGAAACTG GCAAATGCAC CTCTGAGCGG

 401 CAGGTGTTTT TCTCAGGTGT TTTCCGCGTT CCGGGTCAAA GTTGGCGTTT
     GTCCACAAAA AGAGTCCACA AAAGGCGCAA GGCCCAGTTT CAACCGCAAA

 451 TATTATTATA GGCGGCCGCG ATCCATTGCA TACGTTGTAT CCATATCATA
     ATAATAATAT CCGCCGGCGC TAGGTAACGT ATGCAACATA GGTATAGTAT

 501 ATATGTACAT TTATATTGGC TCATGTCCAA CATTACCGCC ATGTTGACAT
     TATACATGTA AATATAACCG AGTACAGGTT GTAATGGCGG TACAACTGTA

 551 TGATTATTGA CTAGTTATTA ATAGTAATCA ATTACGGGGT CATTAGTTCA
     ACTAATAACT GATCAATAAT TATCATTAGT TAATGCCCCA GTAATCAAGT

 601 TAGCCCATAT ATGGAGTTCC GCGTTACATA ACTTACGGTA AATGGCCCGC
     ATCGGGTATA TACCTCAAGG CGCAATGTAT TGAATGCCAT TTACCGGGCG

 651 CTGGCTGACC GCCCAACGAC CCCCGCCCAT TGACGTCAAT AATGACGTAT
     GACCGACTGG CGGGTTGCTG GGGGCGGGTA ACTGCAGTTA TTACTGCATA

 701 GTTCCCATAG TAACGCCAAT AGGGACTTTC CATTGACGTC AATGGGTGGA
     CAAGGGTATC ATTGCGGTTA TCCCTGAAAG GTAACTGCAG TTACCCACCT

 751 GTATTTACGG TAAACTGCCC ACTTGGCAGT ACATCAAGTG TATCATATGC
     CATAAATGCC ATTTGACGGG TGAACCGTCA TGTAGTTCAC ATAGTATACG
```

FIG.9A-1

```
801  CAAGTACGCC CCCTATTGAC GTCAATGACG GTAAATGGCC CGCCTGGCAT
     GTTCATGCGG GGGATAACTG CAGTTACTGC CATTTACCGG GCGGACCGTA

851  TATGCCCAGT ACATGACCTT ATGGGACTTT CCTACTTGGC AGTACATCTA
     ATACGGGTCA TGTACTGGAA TACCCTGAAA GGATGAACCG TCATGTAGAT

901  CGTATTAGTC ATCGCTATTA CCATGGTGAT GCGGTTTTGG CAGTACATCA
     GCATAATCAG TAGCGATAAT GGTACCACTA CGCCAAAACC GTCATGTAGT

951  ATGGGCGTGG ATAGCGGTTT GACTCACGGG GATTTCCAAG TCTCCACCCC
     TACCCGCACC TATCGCCAAA CTGAGTGCCC CTAAAGGTTC AGAGGTGGGG

1001 ATTGACGTCA ATGGGAGTTT GTTTTGGCAC CAAAATCAAC GGGACTTTCC
     TAACTGCAGT TACCCTCAAA CAAAACCGTG GTTTTAGTTG CCCTGAAAGG

1051 AAAATGTCGT AACAACTCCG CCCCATTGAC GCAAATGGGC GGTAGGCGTG
     TTTTACAGCA TTGTTGAGGC GGGGTAACTG CGTTTACCCG CCATCCGCAC

1101 TACGGTGGGA GGTCTATATA AGCAGAGCTC GTTTAGTGAA CCGTCAGATC
     ATGCCACCCT CCAGATATAT TCGTCTCGAG CAAATCACTT GGCAGTCTAG

1151 GCCTGGAGAC GCCATCCACG CTGTTTTGAC CTCCATAGAA GACACCGGGA
     CGGACCTCTG CGGTAGGTGC GACAAAACTG GAGGTATCTT CTGTGGCCCT

1201 CCGATCCAGC CTCCGCGGCC GGGAACGGTG CATTGGAACG CGGATTCCCC
     GGCTAGGTCG GAGGCGCCGG CCCTTGCCAC GTAACCTTGC GCCTAAGGGG

1251 GTGCCAAGAG TGAGATCTAC CATGGGTGCT AGGGCTTCTG TGCTGTCTGG
     CACGGTTCTC ACTCTAGATG GTACCCACGA TCCCGAAGAC ACGACAGACC

1301 TGGTGAGCTG GACAAGTGGG AGAAGATCAG GCTGAGGCCT GGTGGCAAGA
     ACCACTCGAC CTGTTCACCC TCTTCTAGTC CGACTCCGGA CCACCGTTCT

1351 AGAAGTACAA GCTAAAGCAC ATTGTGTGGG CCTCCAGGGA GCTGGAGAGG
     TCTTCATGTT CGATTTCGTG TAACACACCC GGAGGTCCCT CGACCTCTCC

1401 TTTGCTGTGA ACCCTGGCCT GCTGGAGACC TCTGAGGGGT GCAGGCAGAT
     AAACGACACT TGGGACCGGA CGACCTCTGG AGACTCCCCA CGTCCGTCTA

1451 CCTGGGCCAG CTCCAGCCCT CCCTGCAAAC AGGCTCTGAG GAGCTGAGGT
     GGACCCGGTC GAGGTCGGGA GGGACGTTTG TCCGAGACTC CTCGACTCCA

1501 CCCTGTACAA CACAGTGGCT ACCCTGTACT GTGTGCACCA GAAGATTGAT
     GGGACATGTT GTGTCACCGA TGGGACATGA CACACGTGGT CTTCTAACTA

1551 GTGAAGGACA CCAAGGAGGC CCTGGAGAAG ATTGAGGAGG AGCAGAACAA
     CACTTCCTGT GGTTCCTCCG GGACCTCTTC TAACTCCTCC TCGTCTTGTT

1601 GTCCAAGAAG AAGGCCCAGC AGGCTGCTGC TGGCACAGGC AACTCCAGCC
     CAGGTTCTTC TTCCGGGTCG TCCGACGACG ACCGTGTCCG TTGAGGTCGG
```

FIG.9A-2

1651 AGGTGTCCCA GAACTACCCC ATTGTGCAGA ACCTCCAGGG CCAGATGGTG
TCCACAGGGT CTTGATGGGG TAACACGTCT TGGAGGTCCC GGTCTACCAC

1701 CACCAGGCCA TCTCCCCCCG GACCCTGAAT GCCTGGGTGA AGGTGGTGGA
GTGGTCCGGT AGAGGGGGGC CTGGGACTTA CGGACCCACT TCCACCACCT

1751 GGAGAAGGCC TTCTCCCCTG AGGTGATCCC CATGTTCTCT GCCCTGTCTG
CCTCTTCCGG AAGAGGGGAC TCCACTAGGG GTACAAGAGA CGGGACAGAC

1801 AGGGTGCCAC CCCCCAGGAC CTGAACACCA TGCTGAACAC AGTGGGGGGC
TCCCACGGTG GGGGGTCCTG GACTTGTGGT ACGACTTGTG TCACCCCCCG

1851 CATCAGGCTG CCATGCAGAT GCTGAAGGAG ACCATCAATG AGGAGGCTGC
GTAGTCCGAC GGTACGTCTA CGACTTCCTC TGGTAGTTAC TCCTCCGACG

1901 TGAGTGGGAC AGGCTGCATC CTGTGCACGC TGGCCCCATT GCCCCCGGCC
ACTCACCCTG TCCGACGTAG GACACGTGCG ACCGGGGTAA CGGGGGCCGG

1951 AGATGAGGGA GCCCAGGGGC TCTGACATTG CTGGCACCAC CTCCACCCTC
TCTACTCCCT CGGGTCCCCG AGACTGTAAC GACCGTGGTG GAGGTGGGAG

2001 CAGGAGCAGA TTGGCTGGAT GACCAACAAC CCCCCCATCC CTGTGGGGGA
GTCCTCGTCT AACCGACCTA CTGGTTGTTG GGGGGGTAGG GACACCCCCT

2051 AATCTACAAG AGGTGGATCA TCCTGGGCCT GAACAAGATT GTGAGGATGT
TTAGATGTTC TCCACCTAGT AGGACCCGGA CTTGTTCTAA CACTCCTACA

2101 ACTCCCCCAC CTCCATCCTG GACATCAGGC AGGGCCCCAA GGAGCCCTTC
TGAGGGGGTG GAGGTAGGAC CTGTAGTCCG TCCCGGGGTT CCTCGGGAAG

2151 AGGGACTATG TGGACAGGTT CTACAAGACC CTGAGGGCTG AGCAGGCCTC
TCCCTGATAC ACCTGTCCAA GATGTTCTGG GACTCCCGAC TCGTCCGGAG

2201 CCAGGAGGTG AAGAACTGGA TGACAGAGAC CCTGCTGGTG CAGAATGCCA
GGTCCTCCAC TTCTTGACCT ACTGTCTCTG GGACGACCAC GTCTTACGGT

2251 ACCCTGACTG CAAGACCATC CTGAAGGCCC TGGGCCCTGC TGCCACCCTG
TGGGACTGAC GTTCTGGTAG GACTTCCGGG ACCCGGGACG ACGGTGGGAC

2301 GAGGAGATGA TGACAGCCTG CCAGGGGGTG GGGGGCCCTG GTCACAAGGC
CTCCTCTACT ACTGTCGGAC GGTCCCCCAC CCCCCGGGAC CAGTGTTCCG

2351 CAGGGTGCTG GCTGAGGCCA TGTCCCAGGT GACCAACTCC GCCACCATCA
GTCCCACGAC CGACTCCGGT ACAGGGTCCA CTGGTTGAGG CGGTGGTAGT

2401 TGATGCAGAG GGGCAACTTC AGGAACCAGA GGAAGACAGT GAAGTGCTTC
ACTACGTCTC CCCGTTGAAG TCCTTGGTCT CCTTCTGTCA CTTCACGAAG

2451 AACTGTGGCA AGGTGGGCCA CATTGCCAAG AACTGTAGGG CCCCCAGGAA
TTGACACCGT TCCACCCGGT GTAACGGTTC TTGACATCCC GGGGGTCCTT

FIG.9A-3

2501 GAAGGGCTGC TGGAAGTGTG GCAAGGAGGG CCACCAGATG AAGGACTGCA
CTTCCCGACG ACCTTCACAC CGTTCCTCCC GGTGGTCTAC TTCCTGACGT

2551 ATGAGAGGCA GGCCAACTTC CTGGGCAAAA TCTGGCCCTC CCACAAGGGC
TACTCTCCGT CCGGTTGAAG GACCCGTTTT AGACCGGGAG GGTGTTCCCG

2601 AGGCCTGGCA ACTTCCTCCA GTCCAGGCCT GAGCCCACAG CCCCTCCCGA
TCCGGACCGT TGAAGGAGGT CAGGTCCGGA CTCGGGTGTC GGGGAGGGCT

2651 GGAGTCCTTC AGGTTTGGGG AGGAGAAGAC CACCCCCAGC CAGAAGCAGG
CCTCAGGAAG TCCAAACCCC TCCTCTTCTG GTGGGGGTCG GTCTTCGTCC

2701 AGCCCATTGA CAAGGAGCTG TACCCCCTGG CCTCCCTGAG GTCCCTGTTT
TCGGGTAACT GTTCCTCGAC ATGGGGGACC GGAGGGACTC CAGGGACAAA

2751 GGCAACGACC CCTCCTCCCA GTAAAATAAA GCCCGGGCAG ATCTGCTGTG
CCGTTGCTGG GGAGGAGGGT CATTTTATTT CGGGCCCGTC TAGACGACAC

2801 CCTTCTAGTT GCCAGCCATC TGTTGTTTGC CCCTCCCCCG TGCCTTCCTT
GGAAGATCAA CGGTCGGTAG ACAACAAACG GGGAGGGGGC ACGGAAGGAA

2851 GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA AATGAGGAAA
CTGGGACCTT CCACGGTGAG GGTGACAGGA AAGGATTATT TTACTCCTTT

2901 TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG GGGTGGGGTG
AACGTAGCGT AACAGACTCA TCCACAGTAA GATAAGACCC CCCACCCCAC

2951 GGGCAGGACA GCAAGGGGGA GGATTGGGAA GACAATAGCA GGCATGCTGG
CCCGTCCTGT CGTTCCCCCT CCTAACCCTT CTGTTATCGT CCGTACGACC

3001 GGATGCGGTG GGCTCTATGG CCGATCGGCG CGCCGTACTG AAATGTGTGG
CCTACGCCAC CCGAGATACC GGCTAGCCGC GCGGCATGAC TTTACACACC

3051 GCGTGGCTTA AGGGTGGGAA AGAATATATA AGGTGGGGGT CTTATGTAGT
CGCACCGAAT TCCCACCCTT TCTTATATAT TCCACCCCCA GAATACATCA

3101 TTTGTATCTG TTTTGCAGCA GCCGCCGCCG CCATGAGCAC CAACTCGTTT
AAACATAGAC AAAACGTCGT CGGCGGCGGC GGTACTCGTG GTTGAGCAAA

3151 GATGGAAGCA TTGTGAGCTC ATATTTGACA ACGCGCATGC CCCCATGGGC
CTACCTTCGT AACACTCGAG TATAAACTGT TGCGCGTACG GGGGTACCCG

3201 CGGGGTGCGT CAGAATGTGA TGGGCTCCAG CATTGATGGT CGCCCCGTCC
GCCCCACGCA GTCTTACACT ACCCGAGGTC GTAACTACCA GCGGGGCAGG

3251 TGCCCGCAAA CTCTACTACC TTGACCTACG AGACCGTGTC TGGAACGCCG
ACGGGCGTTT GAGATGATGG AACTGGATGC TCTGGCACAG ACCTTGCGGC

3301 TTGGAGACTG CAGCCTCCGC CGCCGCTTCA GCCGCTGCAG CCACCGCCCG
AACCTCTGAC GTCGGAGGCG GCGGCGAAGT CGGCGACGTC GGTGGCGGGC

FIG.9A-4

```
3351  CGGGATTGTG ACTGACTTTG CTTTCCTGAG CCCGCTTGCA AACAGTGCAG
      GCCCTAACAC TGACTGAAAC GAAAGGACTC GGGCGAACGT TTGTCACGTC

3401  CTTCCCGTTC ATCCGCCCGC GATGACAAGT TGACGGCTCT TTTGGCACAA
      GAAGGGCAAG TAGGCGGGCG CTACTGTTCA ACTGCCGAGA AAACCGTGTT

3451  TTGGATTCTT TGACCCGGGA ACTTAATGTC GTTTCTCAGC AGCTGTTGGA
      AACCTAAGAA ACTGGGCCCT TGAATTACAG CAAAGAGTCG TCGACAACCT

3501  TCTGCGCCAG CAGGTTTCTG CCCTGAAGGC TTCCTCCCCT CCCAATGCGG
      AGACGCGGTC GTCCAAAGAC GGGACTTCCG AAGGAGGGGA GGGTTACGCC

3551  TTTAAAACAT AAATAAAAAA CCAGACTCTG TTTGGATTTG GATCAAGCAA
      AAATTTTGTA TTTATTTTTT GGTCTGAGAC AAACCTAAAC CTAGTTCGTT

3601  GTGTCTTGCT GTCTTTATTT AGGGGTTTTG CGCGCGCGGT AGGCCCGGGA
      CACAGAACGA CAGAAATAAA TCCCCAAAAC GCGCGCGCCA TCCGGGCCCT

3651  CCAGCGGTCT CGGTCGTTGA GGGTCCTGTG TATTTTTTCC AGGACGTGGT
      GGTCGCCAGA GCCAGCAACT CCCAGGACAC ATAAAAAAGG TCCTGCACCA

3701  AAAGGTGACT CTGGATGTTC AGATACATGG GCATAAGCCC GTCTCTGGGG
      TTTCCACTGA GACCTACAAG TCTATGTACC CGTATTCGGG CAGAGACCCC

3751  TGGAGGTAGC ACCACTGCAG AGCTTCATGC TGCGGGGTGG TGTTGTAGAT
      ACCTCCATCG TGGTGACGTC TCGAAGTACG ACGCCCCACC ACAACATCTA

3801  GATCCAGTCG TAGCAGGAGC GCTGGGCGTG GTGCCTAAAA ATGTCTTTCA
      CTAGGTCAGC ATCGTCCTCG CGACCCGCAC CACGGATTTT TACAGAAAGT

3851  GTAGCAAGCT GATTGCCAGG GGCAGGCCCT TGGTGTAAGT GTTTACAAAG
      CATCGTTCGA CTAACGGTCC CCGTCCGGGA ACCACATTCA CAAATGTTTC

3901  CGGTTAAGCT GGGATGGGTG CATACGTGGG GATATGAGAT GCATCTTGGA
      GCCAATTCGA CCCTACCCAC GTATGCACCC CTATACTCTA CGTAGAACCT

3951  CTGTATTTTT AGGTTGGCTA TGTTCCCAGC CATATCCCTC CGGGGATTCA
      GACATAAAAA TCCAACCGAT ACAAGGGTCG GTATAGGGAG GCCCCTAAGT

4001  TGTTGTGCAG AACCACCAGC ACAGTGTATC CGGTGCACTT GGGAAATTTG
      ACAACACGTC TTGGTGGTCG TGTCACATAG GCCACGTGAA CCCTTTAAAC

4051  TCATGTAGCT TAGAAGGAAA TGCGTGGAAG AACTTGGAGA CGCCCTTGTG
      AGTACATCGA ATCTTCCTTT ACGCACCTTC TTGAACCTCT GCGGGAACAC

4101  ACCTCCAAGA TTTTCCATGC ATTCGTCCAT AATGATGGCA ATGGGCCCAC
      TGGAGGTTCT AAAAGGTACG TAAGCAGGTA TTACTACCGT TACCCGGGTG

4151  GGGCGGCGGC CTGGGCGAAG ATATTTCTGG GATCACTAAC GTCATAGTTG
      CCCGCCGCCG GACCCGCTTC TATAAAGACC CTAGTGATTG CAGTATCAAC
```

FIG.9A-5

```
4201  TGTTCCAGGA TGAGATCGTC ATAGGCCATT TTTACAAAGC GCGGGCGGAG
      ACAAGGTCCT ACTCTAGCAG TATCCGGTAA AAATGTTTCG CGCCCGCCTC

4251  GGTGCCAGAC TGCGGTATAA TGGTTCCATC CGGCCCAGGG GCGTAGTTAC
      CCACGGTCTG ACGCCATATT ACCAAGGTAG GCCGGGTCCC CGCATCAATG

4301  CCTCACAGAT TTGCATTTCC CACGCTTTGA GTTCAGATGG GGGGATCATG
      GGAGTGTCTA AACGTAAAGG GTGCGAAACT CAAGTCTACC CCCCTAGTAC

4351  TCTACCTGCG GGGCGATGAA GAAAACGGTT TCCGGGGTAG GGGAGATCAG
      AGATGGACGC CCCGCTACTT CTTTTGCCAA AGGCCCCATC CCCTCTAGTC

4401  CTGGGAAGAA AGCAGGTTCC TGAGCAGCTG CGACTTACCG CAGCCGGTGG
      GACCCTTCTT TCGTCCAAGG ACTCGTCGAC GCTGAATGGC GTCGGCCACC

4451  GCCCGTAAAT CACACCTATT ACCGGCTGCA ACTGGTAGTT AAGAGAGCTG
      CGGGCATTTA GTGTGGATAA TGGCCGACGT TGACCATCAA TTCTCTCGAC

4501  CAGCTGCCGT CATCCCTGAG CAGGGGGGCC ACTTCGTTAA GCATGTCCCT
      GTCGACGGCA GTAGGGACTC GTCCCCCCGG TGAAGCAATT CGTACAGGGA

4551  GACTCGCATG TTTTCCCTGA CCAAATCCGC CAGAAGGCGC TCGCCGCCCA
      CTGAGCGTAC AAAAGGGACT GGTTTAGGCG GTCTTCCGCG AGCGGCGGGT

4601  GCGATAGCAG TTCTTGCAAG GAAGCAAAGT TTTTCAACGG TTTGAGACCG
      CGCTATCGTC AAGAACGTTC CTTCGTTTCA AAAAGTTGCC AAACTCTGGC

4651  TCCGCCGTAG GCATGCTTTT GAGCGTTTGA CCAAGCAGTT CCAGGCGGTC
      AGGCGGCATC CGTACGAAAA CTCGCAAACT GGTTCGTCAA GGTCCGCCAG

4701  CCACAGCTCG GTCACCTGCT CTACGGCATC TCGATCCAGC ATATCTCCTC
      GGTGTCGAGC CAGTGGACGA GATGCCGTAG AGCTAGGTCG TATAGAGGAG

4751  GTTTCGCGGG TTGGGGCGGC TTTCGCTGTA CGGCAGTAGT CGGTGCTCGT
      CAAAGCGCCC AACCCCGCCG AAAGCGACAT GCCGTCATCA GCCACGAGCA

4801  CCAGACGGGC CAGGGTCATG TCTTTCCACG GGCGCAGGGT CCTCGTCAGC
      GGTCTGCCCG GTCCCAGTAC AGAAAGGTGC CCGCGTCCCA GGAGCAGTCG

4851  GTAGTCTGGG TCACGGTGAA GGGGTGCGCT CCGGGCTGCG CGCTGGCCAG
      CATCAGACCC AGTGCCACTT CCCCACGCGA GGCCCGACGC GCGACCGGTC

4901  GGTGCGCTTG AGGCTGGTCC TGCTGGTGCT GAAGCGCTGC CGGTCTTCGC
      CCACGCGAAC TCCGACCAGG ACGACCACGA CTTCGCGACG GCCAGAAGCG

4951  CCTGCGCGTC GGCCAGGTAG CATTTGACCA TGGTGTCATA GTCCAGCCCC
      GGACGCGCAG CCGGTCCATC GTAAACTGGT ACCACAGTAT CAGGTCGGGG

5001  TCCGCGGCGT GGCCCTTGGC GCGCAGCTTG CCCTTGGAGG AGGCGCCGCA
      AGGCGCCGCA CCGGGAACCG CGCGTCGAAC GGGAACCTCC TCCGCGGCGT
```

FIG.9A-6

```
5051  CGAGGGGCAG TGCAGACTTT TGAGGGCGTA GAGCTTGGGC GCGAGAAATA
      GCTCCCCGTC ACGTCTGAAA ACTCCCGCAT CTCGAACCCG CGCTCTTTAT

5101  CCGATTCCGG GGAGTAGGCA TCCGCGCCGC AGGCCCCGCA GACGGTCTCG
      GGCTAAGGCC CCTCATCCGT AGGCGCGGCG TCCGGGGCGT CTGCCAGAGC

5151  CATTCCACGA GCCAGGTGAG CTCTGGCCGT TCGGGGTCAA AAACCAGGTT
      GTAAGGTGCT CGGTCCACTC GAGACCGGCA AGCCCCAGTT TTTGGTCCAA

5201  TCCCCCATGC TTTTTGATGC GTTTCTTACC TCTGGTTTCC ATGAGCCGGT
      AGGGGGTACG AAAAACTACG CAAAGAATGG AGACCAAAGG TACTCGGCCA

5251  GTCCACGCTC GGTGACGAAA AGGCTGTCCG TGTCCCCGTA TACAGACTTG
      CAGGTGCGAG CCACTGCTTT TCCGACAGGC ACAGGGGCAT ATGTCTGAAC

5301  AGAGGCCTGT CCTCGAGCGG TGTTCCGCGG TCCTCCTCGT ATAGAAACTC
      TCTCCGGACA GGAGCTCGCC ACAAGGCGCC AGGAGGAGCA TATCTTTGAG

5351  GGACCACTCT GAGACAAAGG CTCGCGTCCA GGCCAGCACG AAGGAGGCTA
      CCTGGTGAGA CTCTGTTTCC GAGCGCAGGT CCGGTCGTGC TTCCTCCGAT

5401  AGTGGGAGGG GTAGCGGTCG TTGTCCACTA GGGGGTCCAC TCGCTCCAGG
      TCACCCTCCC CATCGCCAGC AACAGGTGAT CCCCCAGGTG AGCGAGGTCC

5451  GTGTGAAGAC ACATGTCGCC CTCTTCGGCA TCAAGGAAGG TGATTGGTTT
      CACACTTCTG TGTACAGCGG GAGAAGCCGT AGTTCCTTCC ACTAACCAAA

5501  GTAGGTGTAG GCCACGTGAC CGGGTGTTCC TGAAGGGGGG CTATAAAAGG
      CATCCACATC CGGTGCACTG GCCCACAAGG ACTTCCCCCC GATATTTTCC

5551  GGGTGGGGGC GCGTTCGTCC TCACTCTCTT CCGCATCGCT GTCTGCGAGG
      CCCACCCCCG CGCAAGCAGG AGTGAGAGAA GGCGTAGCGA CAGACGCTCC

5601  GCCAGCTGTT GGGGTGAGTA CTCCCTCTGA AAAGCGGGCA TGACTTCTGC
      CGGTCGACAA CCCCACTCAT GAGGGAGACT TTTCGCCCGT ACTGAAGACG

5651  GCTAAGATTG TCAGTTTCCA AAAACGAGGA GGATTTGATA TTCACCTGGC
      CGATTCTAAC AGTCAAAGGT TTTTGCTCCT CCTAAACTAT AAGTGGACCG

5701  CCGCGGTGAT GCCTTTGAGG GTGGCCGCAT CCATCTGGTC AGAAAAGACA
      GGCGCCACTA CGGAAACTCC CACCGGCGTA GGTAGACCAG TCTTTTCTGT

5751  ATCTTTTTGT TGTCAAGCTT GGTGGCAAAC GACCCGTAGA GGGCGTTGGA
      TAGAAAAACA ACAGTTCGAA CCACCGTTTG CTGGGCATCT CCCGCAACCT

5801  CAGCAACTTG GCGATGGAGC GCAGGGTTTG GTTTTTGTCG CGATCGGCGC
      GTCGTTGAAC CGCTACCTCG CGTCCCAAAC CAAAAACAGC GCTAGCCGCG

5851  GCTCCTTGGC CGCGATGTTT AGCTGCACGT ATTCGCGCGC AACGCACCGC
      CGAGGAACCG GCGCTACAAA TCGACGTGCA TAAGCGCGCG TTGCGTGGCG
```

FIG.9A-7

5901 CATTCGGGAA AGACGGTGGT GCGCTCGTCG GGCACCAGGT GCACGCGCCA
GTAAGCCCTT TCTGCCACCA CGCGAGCAGC CCGTGGTCCA CGTGCGCGGT

5951 ACCGCGGTTG TGCAGGGTGA CAAGGTCAAC GCTGGTGGCT ACCTCTCCGC
TGGCGCCAAC ACGTCCCACT GTTCCAGTTG CGACCACCGA TGGAGAGGCG

6001 GTAGGCGCTC GTTGGTCCAG CAGAGGCGGC CGCCCTTGCG CGAGCAGAAT
CATCCGCGAG CAACCAGGTC GTCTCCGCCG GCGGGAACGC GCTCGTCTTA

6051 GGCGGTAGGG GGTCTAGCTG CGTCTCGTCC GGGGGGTCTG CGTCCACGGT
CCGCCATCCC CCAGATCGAC GCAGAGCAGG CCCCCCAGAC GCAGGTGCCA

6101 AAAGACCCCG GGCAGCAGGC GCGCGTCGAA GTAGTCTATC TTGCATCCTT
TTTCTGGGGC CCGTCGTCCG CGCGCAGCTT CATCAGATAG AACGTAGGAA

6151 GCAAGTCTAG CGCCTGCTGC CATGCGCGGG CGGCAAGCGC GCGCTCGTAT
CGTTCAGATC GCGGACGACG GTACGCGCCC GCCGTTCGCG CGCGAGCATA

6201 GGGTTGAGTG GGGGACCCCA TGGCATGGGG TGGGTGAGCG CGGAGGCGTA
CCCAACTCAC CCCCTGGGGT ACCGTACCCC ACCCACTCGC GCCTCCGCAT

6251 CATGCCGCAA ATGTCGTAAA CGTAGAGGGG CTCTCTGAGT ATTCCAAGAT
GTACGGCGTT TACAGCATTT GCATCTCCCC GAGAGACTCA TAAGGTTCTA

6301 ATGTAGGGTA GCATCTTCCA CCGCGGATGC TGGCGCGCAC GTAATCGTAT
TACATCCCAT CGTAGAAGGT GGCGCCTACG ACCGCGCGTG CATTAGCATA

6351 AGTTCGTGCG AGGGAGCGAG GAGGTCGGGA CCGAGGTTGC TACGGGCGGG
TCAAGCACGC TCCCTCGCTC CTCCAGCCCT GGCTCCAACG ATGCCCGCCC

6401 CTGCTCTGCT CGGAAGACTA TCTGCCTGAA GATGGCATGT GAGTTGGATG
GACGAGACGA GCCTTCTGAT AGACGGACTT CTACCGTACA CTCAACCTAC

6451 ATATGGTTGG ACGCTGGAAG ACGTTGAAGC TGGCGTCTGT GAGACCTACC
TATACCAACC TGCGACCTTC TGCAACTTCG ACCGCAGACA CTCTGGATGG

6501 GCGTCACGCA CGAAGGAGGC GTAGGAGTCG CGCAGCTTGT TGACCAGCTC
CGCAGTGCGT GCTTCCTCCG CATCCTCAGC GCGTCGAACA ACTGGTCGAG

6551 GGCGGTGACC TGCACGTCTA GGGCGCAGTA GTCCAGGGTT TCCTTGATGA
CCGCCACTGG ACGTGCAGAT CCCGCGTCAT CAGGTCCCAA AGGAACTACT

6601 TGTCATACTT ATCCTGTCCC TTTTTTTTCC ACAGCTCGCG GTTGAGGACA
ACAGTATGAA TAGGACAGGG AAAAAAAAGG TGTCGAGCGC CAACTCCTGT

6651 AACTCTTCGC GGTCTTTCCA GTACTCTTGG ATCGGAAACC CGTCGGCCTC
TTGAGAAGCG CCAGAAAGGT CATGAGAACC TAGCCTTTGG GCAGCCGGAG

6701 CGAACGGTAA GAGCCTAGCA TGTAGAACTG GTTGACGGCC TGGTAGGCGC
GCTTGCCATT CTCGGATCGT ACATCTTGAC CAACTGCCGG ACCATCCGCG

FIG.9A-8

6751 AGCATCCCTT TTCTACGGGT AGCGCGTATG CCTGCGCGGC CTTCCGGAGC
TCGTAGGGAA AAGATGCCCA TCGCGCATAC GGACGCGCCG GAAGGCCTCG

6801 GAGGTGTGGG TGAGCGCAAA GGTGTCCCTG ACCATGACTT TGAGGTACTG
CTCCACACCC ACTCGCGTTT CCACAGGGAC TGGTACTGAA ACTCCATGAC

6851 GTATTTGAAG TCAGTGTCGT CGCATCCGCC CTGCTCCCAG AGCAAAAAGT
CATAAACTTC AGTCACAGCA GCGTAGGCGG GACGAGGGTC TCGTTTTTCA

6901 CCGTGCGCTT TTTGGAACGC GGATTTGGCA GGGCGAAGGT GACATCGTTG
GGCACGCGAA AAACCTTGCG CCTAAACCGT CCCGCTTCCA CTGTAGCAAC

6951 AAGAGTATCT TTCCCGCGCG AGGCATAAAG TTGCGTGTGA TGCGGAAGGG
TTCTCATAGA AAGGGCGCGC TCCGTATTTC AACGCACACT ACGCCTTCCC

7001 TCCCGGCACC TCGGAACGGT TGTTAATTAC CTGGGCGGCG AGCACGATCT
AGGGCCGTGG AGCCTTGCCA ACAATTAATG GACCCGCCGC TCGTGCTAGA

7051 CGTCAAAGCC GTTGATGTTG TGGCCCACAA TGTAAAGTTC CAAGAAGCGC
GCAGTTTCGG CAACTACAAC ACCGGGTGTT ACATTTCAAG GTTCTTCGCG

7101 GGGATGCCCT TGATGGAAGG CAATTTTTTA AGTTCCTCGT AGGTGAGCTC
CCCTACGGGA ACTACCTTCC GTTAAAAAAT TCAAGGAGCA TCCACTCGAG

7151 TTCAGGGGAG CTGAGCCCGT GCTCTGAAAG GGCCCAGTCT GCAAGATGAG
AAGTCCCCTC GACTCGGGCA CGAGACTTTC CCGGGTCAGA CGTTCTACTC

7201 GGTTGGAAGC GACGAATGAG CTCCACAGGT CACGGGCCAT TAGCATTTGC
CCAACCTTCG CTGCTTACTC GAGGTGTCCA GTGCCCGGTA ATCGTAAACG

7251 AGGTGGTCGC GAAAGGTCCT AAACTGGCGA CCTATGGCCA TTTTTTCTGG
TCCACCAGCG CTTTCCAGGA TTTGACCGCT GGATACCGGT AAAAAAGACC

7301 GGTGATGCAG TAGAAGGTAA GCGGGTCTTG TTCCCAGCGG TCCCATCCAA
CCACTACGTC ATCTTCCATT CGCCCAGAAC AAGGGTCGCC AGGGTAGGTT

7351 GGTTCGCGGC TAGGTCTCGC GCGGCAGTCA CTAGAGGCTC ATCTCCGCCG
CCAAGCGCCG ATCCAGAGCG CGCCGTCAGT GATCTCCGAG TAGAGGCGGC

7401 AACTTCATGA CCAGCATGAA GGGCACGAGC TGCTTCCCAA AGGCCCCCAT
TTGAAGTACT GGTCGTACTT CCCGTGCTCG ACGAAGGGTT TCCGGGGGTA

7451 CCAAGTATAG GTCTCTACAT CGTAGGTGAC AAAGAGACGC TCGGTGCGAG
GGTTCATATC CAGAGATGTA GCATCCACTG TTTCTCTGCG AGCCACGCTC

7501 GATGCGAGCC GATCGGGAAG AACTGGATCT CCCGCCACCA ATTGGAGGAG
CTACGCTCGG CTAGCCCTTC TTGACCTAGA GGGCGGTGGT TAACCTCCTC

7551 TGGCTATTGA TGTGGTGAAA GTAGAAGTCC CTGCGACGGG CCGAACACTC
ACCGATAACT ACACCACTTT CATCTTCAGG GACGCTGCCC GGCTTGTGAG

FIG.9A-9

7601 GTGCTGGCTT TTGTAAAAAC GTGCGCAGTA CTGGCAGCGG TGCACGGGCT
CACGACCGAA AACATTTTTG CACGCGTCAT GACCGTCGCC ACGTGCCCGA

7651 GTACATCCTG CACGAGGTTG ACCTGACGAC CGCGCACAAG GAAGCAGAGT
CATGTAGGAC GTGCTCCAAC TGGACTGCTG GCGCGTGTTC CTTCGTCTCA

7701 GGGAATTTGA GCCCCTCGCC TGGCGGGTTT GGCTGGTGGT CTTCTACTTC
CCCTTAAACT CGGGGAGCGG ACCGCCCAAA CCGACCACCA GAAGATGAAG

7751 GGCTGCTTGT CCTTGACCGT CTGGCTGCTC GAGGGGAGTT ACGGTGGATC
CCGACGAACA GGAACTGGCA GACCGACGAG CTCCCCTCAA TGCCACCTAG

7801 GGACCACCAC GCCGCGCGAG CCCAAAGTCC AGATGTCCGC GCGCGGCGGT
CCTGGTGGTG CGGCGCGCTC GGGTTTCAGG TCTACAGGCG CGCGCCGCCA

7851 CGGAGCTTGA TGACAACATC GCGCAGATGG GAGCTGTCCA TGGTCTGGAG
GCCTCGAACT ACTGTTGTAG CGCGTCTACC CTCGACAGGT ACCAGACCTC

7901 CTCCCGCGGC GTCAGGTCAG GCGGGAGCTC CTGCAGGTTT ACCTCGCATA
GAGGGCGCCG CAGTCCAGTC CGCCCTCGAG GACGTCCAAA TGGAGCGTAT

7951 GACGGGTCAG GGCGCGGGCT AGATCCAGGT GATACCTAAT TTCCAGGGGC
CTGCCCAGTC CCGCGCCCGA TCTAGGTCCA CTATGGATTA AAGGTCCCCG

8001 TGGTTGGTGG CGGCGTCGAT GGCTTGCAAG AGGCCGCATC CCCGCGGCGC
ACCAACCACC GCCGCAGCTA CCGAACGTTC TCCGGCGTAG GGGCGCCGCG

8051 GACTACGGTA CCGCGCGGCG GGCGGTGGGC CGCGGGGGTG TCCTTGGATG
CTGATGCCAT GGCGCGCCGC CCGCCACCCG GCGCCCCCAC AGGAACCTAC

8101 ATGCATCTAA AAGCGGTGAC GCGGGCGAGC CCCCGGAGGT AGGGGGGGCT
TACGTAGATT TTCGCCACTG CGCCCGCTCG GGGGCCTCCA TCCCCCCCGA

8151 CCGGACCCGC CGGGAGAGGG GGCAGGGGCA CGTCGGCGCC GCGCGCGGGC
GGCCTGGGCG GCCCTCTCCC CCGTCCCCGT GCAGCCGCGG CGCGCGCCCG

8201 AGGAGCTGGT GCTGCGCGCG TAGGTTGCTG GCGAACGCGA CGACGCGGCG
TCCTCGACCA CGACGCGCGC ATCCAACGAC CGCTTGCGCT GCTGCGCCGC

8251 GTTGATCTCC TGAATCTGGC GCCTCTGCGT GAAGACGACG GGCCCGGTGA
CAACTAGAGG ACTTAGACCG CGGAGACGCA CTTCTGCTGC CCGGGCCACT

8301 GCTTGAACCT GAAAGAGAGT TCGACAGAAT CAATTTCGGT GTCGTTGACG
CGAACTTGGA CTTTCTCTCA AGCTGTCTTA GTTAAAGCCA CAGCAACTGC

8351 GCGGCCTGGC GCAAAATCTC CTGCACGTCT CCTGAGTTGT CTTGATAGGC
CGCCGGACCG CGTTTTAGAG GACGTGCAGA GGACTCAACA GAACTATCCG

8401 GATCTCGGCC ATGAACTGCT CGATCTCTTC CTCCTGGAGA TCTCCGCGTC
CTAGAGCCGG TACTTGACGA GCTAGAGAAG GAGGACCTCT AGAGGCGCAG

FIG.9A-10

8451 CGGCTCGCTC CACGGTGGCG GCGAGGTCGT TGGAAATGCG GGCCATGAGC
GCCGAGCGAG GTGCCACCGC CGCTCCAGCA ACCTTTACGC CCGGTACTCG

8501 TGCGAGAAGG CGTTGAGGCC TCCCTCGTTC CAGACGCGGC TGTAGACCAC
ACGCTCTTCC GCAACTCCGG AGGGAGCAAG GTCTGCGCCG ACATCTGGTG

8551 GCCCCCTTCG GCATCGCGGG CGCGCATGAC CACCTGCGCG AGATTGAGCT
CGGGGGAAGC CGTAGCGCCC GCGCGTACTG GTGGACGCGC TCTAACTCGA

8601 CCACGTGCCG GGCGAAGACG GCGTAGTTTC GCAGGCGCTG AAAGAGGTAG
GGTGCACGGC CCGCTTCTGC CGCATCAAAG CGTCCGCGAC TTTCTCCATC

8651 TTGAGGGTGG TGGCGGTGTG TTCTGCCACG AAGAAGTACA TAACCCAGCG
AACTCCCACC ACCGCCACAC AAGACGGTGC TTCTTCATGT ATTGGGTCGC

8701 TCGCAACGTG GATTCGTTGA TATCCCCCAA GGCCTCAAGG CGCTCCATGG
AGCGTTGCAC CTAAGCAACT ATAGGGGGTT CCGGAGTTCC GCGAGGTACC

8751 CCTCGTAGAA GTCCACGGCG AAGTTGAAAA ACTGGGAGTT GCGCGCCGAC
GGAGCATCTT CAGGTGCCGC TTCAACTTTT TGACCCTCAA CGCGCGGCTG

8801 ACGGTTAACT CCTCCTCCAG AAGACGGATG AGCTCGGCGA CAGTGTCGCG
TGCCAATTGA GGAGGAGGTC TTCTGCCTAC TCGAGCCGCT GTCACAGCGC

8851 CACCTCGCGC TCAAAGGCTA CAGGGGCCTC TTCTTCTTCT TCAATCTCCT
GTGGAGCGCG AGTTTCCGAT GTCCCCGGAG AAGAAGAAGA AGTTAGAGGA

8901 CTTCCATAAG GGCCTCCCCT TCTTCTTCTT CTGGCGGCGG TGGGGGAGGG
GAAGGTATTC CCGGAGGGGA AGAAGAAGAA GACCGCCGCC ACCCCCTCCC

8951 GGGACACGGC GGCGACGACG GCGCACCGGG AGGCGGTCGA CAAAGCGCTC
CCCTGTGCCG CCGCTGCTGC CGCGTGGCCC TCCGCCAGCT GTTTCGCGAG

9001 GATCATCTCC CCGCGGCGAC GGCGCATGGT CTCGGTGACG GCGCGGCCGT
CTAGTAGAGG GGCGCCGCTG CCGCGTACCA GAGCCACTGC CGCGCCGGCA

9051 TCTCGCGGGG GCGCAGTTGG AAGACGCCGC CCGTCATGTC CCGGTTATGG
AGAGCGCCCC CGCGTCAACC TTCTGCGGCG GGCAGTACAG GGCCAATACC

9101 GTTGGCGGGG GGCTGCCATG CGGCAGGGAT ACGGCGCTAA CGATGCATCT
CAACCGCCCC CCGACGGTAC GCCGTCCCTA TGCCGCGATT GCTACGTAGA

9151 CAACAATTGT TGTGTAGGTA CTCCGCCGCC GAGGGACCTG AGCGAGTCCG
GTTGTTAACA ACACATCCAT GAGGCGGCGG CTCCCTGGAC TCGCTCAGGC

9201 CATCGACCGG ATCGGAAAAC CTCTCGAGAA AGGCGTCTAA CCAGTCACAG
GTAGCTGGCC TAGCCTTTTG GAGAGCTCTT TCCGCAGATT GGTCAGTGTC

9251 TCGCAAGGTA GGCTGAGCAC CGTGGCGGGC GGCAGCGGGC GGCGGTCGGG
AGCGTTCCAT CCGACTCGTG GCACCGCCCG CCGTCGCCCG CCGCCAGCCC

FIG.9A-11

9301 GTTGTTTCTG GCGGAGGTGC TGCTGATGAT GTAATTAAAG TAGGCGGTCT
CAACAAAGAC CGCCTCCACG ACGACTACTA CATTAATTTC ATCCGCCAGA

9351 TGAGACGGCG GATGGTCGAC AGAAGCACCA TGTCCTTGGG TCCGGCCTGC
ACTCTGCCGC CTACCAGCTG TCTTCGTGGT ACAGGAACCC AGGCCGGACG

9401 TGAATGCGCA GGCGGTCGGC CATGCCCCAG GCTTCGTTTT GACATCGGCG
ACTTACGCGT CCGCCAGCCG GTACGGGGTC CGAAGCAAAA CTGTAGCCGC

9451 CAGGTCTTTG TAGTAGTCTT GCATGAGCCT TTCTACCGGC ACTTCTTCTT
GTCCAGAAAC ATCATCAGAA CGTACTCGGA AAGATGGCCG TGAAGAAGAA

9501 CTCCTTCCTC TTGTCCTGCA TCTCTTGCAT CTATCGCTGC GGCGGCGGCG
GAGGAAGGAG AACAGGACGT AGAGAACGTA GATAGCGACG CCGCCGCCGC

9551 GAGTTTGGCC GTAGGTGGCG CCCTCTTCCT CCCATGCGTG TGACCCCGAA
CTCAAACCGG CATCCACCGC GGGAGAAGGA GGGTACGCAC ACTGGGGCTT

9601 GCCCCTCATC GGCTGAAGCA GGGCTAGGTC GGCGACAACG CGCTCGGCTA
CGGGGAGTAG CCGACTTCGT CCCGATCCAG CCGCTGTTGC GCGAGCCGAT

9651 ATATGGCCTG CTGCACCTGC GTGAGGGTAG ACTGGAAGTC ATCCATGTCC
TATACCGGAC GACGTGGACG CACTCCCATC TGACCTTCAG TAGGTACAGG

9701 ACAAAGCGGT GGTATGCGCC CGTGTTGATG GTGTAAGTGC AGTTGGCCAT
TGTTTCGCCA CCATACGCGG GCACAACTAC CACATTCACG TCAACCGGTA

9751 AACGGACCAG TTAACGGTCT GGTGACCCGG CTGCGAGAGC TCGGTGTACC
TTGCCTGGTC AATTGCCAGA CCACTGGGCC GACGCTCTCG AGCCACATGG

9801 TGAGACGCGA GTAAGCCCTC GAGTCAAATA CGTAGTCGTT GCAAGTCCGC
ACTCTGCGCT CATTCGGGAG CTCAGTTTAT GCATCAGCAA CGTTCAGGCG

9851 ACCAGGTACT GGTATCCCAC CAAAAAGTGC GGCGGCGGCT GGCGGTAGAG
TGGTCCATGA CCATAGGGTG GTTTTTCACG CCGCCGCCGA CCGCCATCTC

9901 GGGCCAGCGT AGGGTGGCCG GGGCTCCGGG GGCGAGATCT TCCAACATAA
CCCGGTCGCA TCCCACCGGC CCCGAGGCCC CCGCTCTAGA AGGTTGTATT

9951 GGCGATGATA TCCGTAGATG TACCTGGACA TCCAGGTGAT GCCGGCGGCG
CCGCTACTAT AGGCATCTAC ATGGACCTGT AGGTCCACTA CGGCCGCCGC

10001 GTGGTGGAGG CGCGCGGAAA GTCGCGGACG CGGTTCCAGA TGTTGCGCAG
CACCACCTCC GCGCGCCTTT CAGCGCCTGC GCCAAGGTCT ACAACGCGTC

10051 CGGCAAAAAG TGCTCCATGG TCGGGACGCT CTGGCCGGTC AGGCGCGCGC
GCCGTTTTTC ACGAGGTACC AGCCCTGCGA GACCGGCCAG TCCGCGCGCG

10101 AATCGTTGAC GCTCTAGACC GTGCAAAAGG AGAGCCTGTA AGCGGGCACT
TTAGCAACTG CGAGATCTGG CACGTTTTCC TCTCGGACAT TCGCCCGTGA

FIG.9A-12

10151 CTTCCGTGGT CTGGTGGATA AATTCGCAAG GGTATCATGG CGGACGACCG
GAAGGCACCA GACCACCTAT TTAAGCGTTC CCATAGTACC GCCTGCTGGC

10201 GGGTTCGAGC CCCGTATCCG GCCGTCCGCC GTGATCCATG CGGTTACCGC
CCCAAGCTCG GGGCATAGGC CGGCAGGCGG CACTAGGTAC GCCAATGGCG

10251 CCGCGTGTCG AACCCAGGTG TGCGACGTCA GACAACGGGG GAGTGCTCCT
GGCGCACAGC TTGGGTCCAC ACGCTGCAGT CTGTTGCCCC CTCACGAGGA

10301 TTTGGCTTCC TTCCAGGCGC GGCGGCTGCT GCGCTAGCTT TTTTGGCCAC
AAACCGAAGG AAGGTCCGCG CCGCCGACGA CGCGATCGAA AAAACCGGTG

10351 TGGCCGCGCG CAGCGTAAGC GGTTAGGCTG GAAAGCGAAA GCATTAAGTG
ACCGGCGCGC GTCGCATTCG CCAATCCGAC CTTTCGCTTT CGTAATTCAC

10401 GCTCGCTCCC TGTAGCCGGA GGGTTATTTT CCAAGGGTTG AGTCGCGGGA
CGAGCGAGGG ACATCGGCCT CCCAATAAAA GGTTCCCAAC TCAGCGCCCT

10451 CCCCCGGTTC GAGTCTCGGA CCGGCCGGAC TGCGGCGAAC GGGGGTTTGC
GGGGGCCAAG CTCAGAGCCT GGCCGGCCTG ACGCCGCTTG CCCCCAAACG

10501 CTCCCCGTCA TGCAAGACCC CGCTTGCAAA TTCCTCCGGA AACAGGGACG
GAGGGGCAGT ACGTTCTGGG GCGAACGTTT AAGGAGGCCT TTGTCCCTGC

10551 AGCCCCTTTT TTGCTTTTCC CAGATGCATC CGGTGCTGCG GCAGATGCGC
TCGGGGAAAA AACGAAAAGG GTCTACGTAG GCCACGACGC CGTCTACGCG

10601 CCCCCTCCTC AGCAGCGGCA AGAGCAAGAG CAGCGGCAGA CATGCAGGGC
GGGGGAGGAG TCGTCGCCGT TCTCGTTCTC GTCGCCGTCT GTACGTCCCG

10651 ACCCTCCCCT CCTCCTACCG CGTCAGGAGG GGCGACATCC GCGGTTGACG
TGGGAGGGGA GGAGGATGGC GCAGTCCTCC CCGCTGTAGG CGCCAACTGC

10701 CGGCAGCAGA TGGTGATTAC GAACCCCCGC GGCGCCGGGC CCGGCACTAC
GCCGTCGTCT ACCACTAATG CTTGGGGGCG CCGCGGCCCG GGCCGTGATG

10751 CTGGACTTGG AGGAGGGCGA GGGCCTGGCG CGGCTAGGAG CGCCCTCTCC
GACCTGAACC TCCTCCCGCT CCCGGACCGC GCCGATCCTC GCGGGAGAGG

10801 TGAGCGGCAC CCAAGGGTGC AGCTGAAGCG TGATACGCGT GAGGCGTACG
ACTCGCCGTG GGTTCCCACG TCGACTTCGC ACTATGCGCA CTCCGCATGC

10851 TGCCGCGGCA GAACCTGTTT CGCGACCGCG AGGGAGAGGA GCCCGAGGAG
ACGGCGCCGT CTTGGACAAA GCGCTGGCGC TCCCTCTCCT CGGGCTCCTC

10901 ATGCGGGATC GAAAGTTCCA CGCAGGGCGC GAGCTGCGGC ATGGCCTGAA
TACGCCCTAG CTTTCAAGGT GCGTCCCGCG CTCGACGCCG TACCGGACTT

10951 TCGCGAGCGG TTGCTGCGCG AGGAGGACTT TGAGCCCGAC GCGCGAACCG
AGCGCTCGCC AACGACGCGC TCCTCCTGAA ACTCGGGCTG CGCGCTTGGC

FIG.9A-13

11001 GGATTAGTCC CGCGCGCGCA CACGTGGCGG CCGCCGACCT GGTAACCGCA
CCTAATCAGG GCGCGCGCGT GTGCACCGCC GGCGGCTGGA CCATTGGCGT

11051 TACGAGCAGA CGGTGAACCA GGAGATTAAC TTTCAAAAAA GCTTTAACAA
ATGCTCGTCT GCCACTTGGT CCTCTAATTG AAAGTTTTTT CGAAATTGTT

11101 CCACGTGCGT ACGCTTGTGG CGCGCGAGGA GGTGGCTATA GGACTGATGC
GGTGCACGCA TGCGAACACC GCGCGCTCCT CCACCGATAT CCTGACTACG

11151 ATCTGTGGGA CTTTGTAAGC GCGCTGGAGC AAAACCCAAA TAGCAAGCCG
TAGACACCCT GAAACATTCG CGCGACCTCG TTTTGGGTTT ATCGTTCGGC

11201 CTCATGGCGC AGCTGTTCCT TATAGTGCAG CACAGCAGGG ACAACGAGGC
GAGTACCGCG TCGACAAGGA ATATCACGTC GTGTCGTCCC TGTTGCTCCG

11251 ATTCAGGGAT GCGCTGCTAA ACATAGTAGA GCCCGAGGGC CGCTGGCTGC
TAAGTCCCTA CGCGACGATT TGTATCATCT CGGGCTCCCG GCGACCGACG

11301 TCGATTTGAT AAACATCCTG CAGAGCATAG TGGTGCAGGA GCGCAGCTTG
AGCTAAACTA TTTGTAGGAC GTCTCGTATC ACCACGTCCT CGCGTCGAAC

11351 AGCCTGGCTG ACAAGGTGGC CGCCATCAAC TATTCCATGC TTAGCCTGGG
TCGGACCGAC TGTTCCACCG GCGGTAGTTG ATAAGGTACG AATCGGACCC

11401 CAAGTTTTAC GCCCGCAAGA TATACCATAC CCCTTACGTT CCCATAGACA
GTTCAAAATG CGGGCGTTCT ATATGGTATG GGGAATGCAA GGGTATCTGT

11451 AGGAGGTAAA GATCGAGGGG TTCTACATGC GCATGGCGCT GAAGGTGCTT
TCCTCCATTT CTAGCTCCCC AAGATGTACG CGTACCGCGA CTTCCACGAA

11501 ACCTTGAGCG ACGACCTGGG CGTTTATCGC AACGAGCGCA TCCACAAGGC
TGGAACTCGC TGCTGGACCC GCAAATAGCG TTGCTCGCGT AGGTGTTCCG

11551 CGTGAGCGTG AGCCGGCGGC GCGAGCTCAG CGACCGCGAG CTGATGCACA
GCACTCGCAC TCGGCCGCCG CGCTCGAGTC GCTGGCGCTC GACTACGTGT

11601 GCCTGCAAAG GGCCCTGGCT GGCACGGGCA GCGGCGATAG AGAGGCCGAG
CGGACGTTTC CCGGGACCGA CCGTGCCCGT CGCCGCTATC TCTCCGGCTC

11651 TCCTACTTTG ACGCGGGCGC TGACCTGCGC TGGGCCCCAA GCCGACGCGC
AGGATGAAAC TGCGCCCGCG ACTGGACGCG ACCCGGGGTT CGGCTGCGCG

11701 CCTGGAGGCA GCTGGGGCCG GACCTGGGCT GGCGGTGGCA CCCGCGCGCG
GGACCTCCGT CGACCCCGGC CTGGACCCGA CCGCCACCGT GGGCGCGCGC

11751 CTGGCAACGT CGGCGGCGTG GAGGAATATG ACGAGGACGA TGAGTACGAG
GACCGTTGCA GCCGCCGCAC CTCCTTATAC TGCTCCTGCT ACTCATGCTC

11801 CCAGAGGACG GCGAGTACTA AGCGGTGATG TTTCTGATCA GATGATGCAA
GGTCTCCTGC CGCTCATGAT TCGCCACTAC AAAGACTAGT CTACTACGTT

FIG.9A-14

```
11851  GACGCAACGG ACCCGGCGGT GCGGGCGGCG CTGCAGAGCC AGCCGTCCGG
       CTGCGTTGCC TGGGCCGCCA CGCCCGCCGC GACGTCTCGG TCGGCAGGCC

11901  CCTTAACTCC ACGGACGACT GGCGCCAGGT CATGGACCGC ATCATGTCGC
       GGAATTGAGG TGCCTGCTGA CCGCGGTCCA GTACCTGGCG TAGTACAGCG

11951  TGACTGCGCG CAATCCTGAC GCGTTCCGGC AGCAGCCGCA GGCCAACCGG
       ACTGACGCGC GTTAGGACTG CGCAAGGCCG TCGTCGGCGT CCGGTTGGCC

12001  CTCTCCGCAA TTCTGGAAGC GGTGGTCCCG GCGCGCGCAA ACCCCACGCA
       GAGAGGCGTT AAGACCTTCG CCACCAGGGC CGCGCGCGTT TGGGGTGCGT

12051  CGAGAAGGTG CTGGCGATCG TAAACGCGCT GGCCGAAAAC AGGGCCATCC
       GCTCTTCCAC GACCGCTAGC ATTTGCGCGA CCGGCTTTTG TCCCGGTAGG

12101  GGCCCGACGA GGCCGGCCTG GTCTACGACG CGCTGCTTCA GCGCGTGGCT
       CCGGGCTGCT CCGGCCGGAC CAGATGCTGC GCGACGAAGT CGCGCACCGA

12151  CGTTACAACA GCGGCAACGT GCAGACCAAC CTGGACCGGC TGGTGGGGGA
       GCAATGTTGT CGCCGTTGCA CGTCTGGTTG GACCTGGCCG ACCACCCCCT

12201  TGTGCGCGAG GCCGTGGCGC AGCGTGAGCG CGCGCAGCAG CAGGGCAACC
       ACACGCGCTC CGGCACCGCG TCGCACTCGC GCGCGTCGTC GTCCCGTTGG

12251  TGGGCTCCAT GGTTGCACTA AACGCCTTCC TGAGTACACA GCCCGCCAAC
       ACCCGAGGTA CCAACGTGAT TTGCGGAAGG ACTCATGTGT CGGGCGGTTG

12301  GTGCCGCGGG GACAGGAGGA CTACACCAAC TTTGTGAGCG CACTGCGGCT
       CACGGCGCCC CTGTCCTCCT GATGTGGTTG AAACACTCGC GTGACGCCGA

12351  AATGGTGACT GAGACACCGC AAAGTGAGGT GTACCAGTCT GGGCCAGACT
       TTACCACTGA CTCTGTGGCG TTTCACTCCA CATGGTCAGA CCCGGTCTGA

12401  ATTTTTTCCA GACCAGTAGA CAAGGCCTGC AGACCGTAAA CCTGAGCCAG
       TAAAAAAGGT CTGGTCATCT GTTCCGGACG TCTGGCATTT GGACTCGGTC

12451  GCTTTCAAAA ACTTGCAGGG GCTGTGGGGG GTGCGGGCTC CCACAGGCGA
       CGAAAGTTTT TGAACGTCCC CGACACCCCC CACGCCCGAG GGTGTCCGCT

12501  CCGCGCGACC GTGTCTAGCT TGCTGACGCC CAACTCGCGC CTGTTGCTGC
       GGCGCGCTGG CACAGATCGA ACGACTGCGG GTTGAGCGCG GACAACGACG

12551  TGCTAATAGC GCCCTTCACG GACAGTGGCA GCGTGTCCCG GGACACATAC
       ACGATTATCG CGGGAAGTGC CTGTCACCGT CGCACAGGGC CCTGTGTATG

12601  CTAGGTCACT TGCTGACACT GTACCGCGAG GCCATAGGTC AGGCGCATGT
       GATCCAGTGA ACGACTGTGA CATGGCGCTC CGGTATCCAG TCCGCGTACA

12651  GGACGAGCAT ACTTTCCAGG AGATTACAAG TGTCAGCCGC GCGCTGGGGC
       CCTGCTCGTA TGAAAGGTCC TCTAATGTTC ACAGTCGGCG CGCGACCCCG
```

FIG.9A-15

12701 AGGAGGACAC GGGCAGCCTG GAGGCAACCC TAAACTACCT GCTGACCAAC
TCCTCCTGTG CCCGTCGGAC CTCCGTTGGG ATTTGATGGA CGACTGGTTG

12751 CGGCGGCAGA AGATCCCCTC GTTGCACAGT TTAAACAGCG AGGAGGAGCG
GCCGCCGTCT TCTAGGGGAG CAACGTGTCA AATTTGTCGC TCCTCCTCGC

12801 CATTTTGCGC TACGTGCAGC AGAGCGTGAG CCTTAACCTG ATGCGCGACG
GTAAAACGCG ATGCACGTCG TCTCGCACTC GGAATTGGAC TACGCGCTGC

12851 GGGTAACGCC CAGCGTGGCG CTGGACATGA CCGCGCGCAA CATGGAACCG
CCCATTGCGG GTCGCACCGC GACCTGTACT GGCGCGCGTT GTACCTTGGC

12901 GGCATGTATG CCTCAAACCG GCCGTTTATC AACCGCCTAA TGGACTACTT
CCGTACATAC GGAGTTTGGC CGGCAAATAG TTGGCGGATT ACCTGATGAA

12951 GCATCGCGCG GCCGCCGTGA ACCCCGAGTA TTTCACCAAT GCCATCTTGA
CGTAGCGCGC CGGCGGCACT TGGGGCTCAT AAAGTGGTTA CGGTAGAACT

13001 ACCCGCACTG GCTACCGCCC CCTGGTTTCT ACACCGGGGG ATTCGAGGTG
TGGGCGTGAC CGATGGCGGG GGACCAAAGA TGTGGCCCCC TAAGCTCCAC

13051 CCCGAGGGTA ACGATGGATT CCTCTGGGAC GACATAGACG ACAGCGTGTT
GGGCTCCCAT TGCTACCTAA GGAGACCCTG CTGTATCTGC TGTCGCACAA

13101 TTCCCCGCAA CCGCAGACCC TGCTAGAGTT GCAACAGCGC GAGCAGGCAG
AAGGGGCGTT GGCGTCTGGG ACGATCTCAA CGTTGTCGCG CTCGTCCGTC

13151 AGGCGGCGCT GCGAAAGGAA AGCTTCCGCA GGCCAAGCAG CTTGTCCGAT
TCCGCCGCGA CGCTTTCCTT TCGAAGGCGT CCGGTTCGTC GAACAGGCTA

13201 CTAGGCGCTG CGGCCCCGCG GTCAGATGCT AGTAGCCCAT TTCCAAGCTT
GATCCGCGAC GCCGGGGCGC CAGTCTACGA TCATCGGGTA AAGGTTCGAA

13251 GATAGGGTCT CTTACCAGCA CTCGCACCAC CCGCCCGCGC CTGCTGGGCG
CTATCCCAGA GAATGGTCGT GAGCGTGGTG GGCGGGCGCG GACGACCCGC

13301 AGGAGGAGTA CCTAAACAAC TCGCTGCTGC AGCCGCAGCG CGAAAAAAAC
TCCTCCTCAT GGATTTGTTG AGCGACGACG TCGGCGTCGC GCTTTTTTTG

13351 CTGCCTCCGG CATTTCCCAA CAACGGGATA GAGAGCCTAG TGGACAAGAT
GACGGAGGCC GTAAAGGGTT GTTGCCCTAT CTCTCGGATC ACCTGTTCTA

13401 GAGTAGATGG AAGACGTACG CGCAGGAGCA CAGGGACGTG CCAGGCCCGC
CTCATCTACC TTCTGCATGC GCGTCCTCGT GTCCCTGCAC GGTCCGGGCG

13451 GCCCGCCCAC CCGTCGTCAA AGGCACGACC GTCAGCGGGG TCTGGTGTGG
CGGGCGGGTG GGCAGCAGTT TCCGTGCTGG CAGTCGCCCC AGACCACACC

13501 GAGGACGATG ACTCGGCAGA CGACAGCAGC GTCCTGGATT TGGGAGGGAG
CTCCTGCTAC TGAGCCGTCT GCTGTCGTCG CAGGACCTAA ACCCTCCCTC

FIG.9A-16

13551 TGGCAACCCG TTTGCGCACC TTCGCCCCAG GCTGGGGAGA ATGTTTTAAA
ACCGTTGGGC AAACGCGTGG AAGCGGGGTC CGACCCCTCT TACAAAATTT

13601 AAAAAAAAAA GCATGATGCA AAATAAAAAA CTCACCAAGG CCATGGCACC
TTTTTTTTTT CGTACTACGT TTTATTTTTT GAGTGGTTCC GGTACCGTGG

13651 GAGCGTTGGT TTTCTTGTAT TCCCCTTAGT ATGCGGCGCG CGGCGATGTA
CTCGCAACCA AAAGAACATA AGGGGAATCA TACGCCGCGC GCCGCTACAT

13701 TGAGGAAGGT CCTCCTCCCT CCTACGAGAG TGTGGTGAGC GCGGCGCCAG
ACTCCTTCCA GGAGGAGGGA GGATGCTCTC ACACCACTCG CGCCGCGGTC

13751 TGGCGGCGGC GCTGGGTTCT CCCTTCGATG CTCCCCTGGA CCCGCCGTTT
ACCGCCGCCG CGACCCAAGA GGGAAGCTAC GAGGGGACCT GGGCGGCAAA

13801 GTGCCTCCGC GGTACCTGCG GCCTACCGGG GGGAGAAACA GCATCCGTTA
CACGGAGGCG CCATGGACGC CGGATGGCCC CCCTCTTTGT CGTAGGCAAT

13851 CTCTGAGTTG GCACCCCTAT TCGACACCAC CCGTGTGTAC CTGGTGGACA
GAGACTCAAC CGTGGGGATA AGCTGTGGTG GGCACACATG GACCACCTGT

13901 ACAAGTCAAC GGATGTGGCA TCCCTGAACT ACCAGAACGA CCACAGCAAC
TGTTCAGTTG CCTACACCGT AGGGACTTGA TGGTCTTGCT GGTGTCGTTG

13951 TTTCTGACCA CGGTCATTCA AAACAATGAC TACAGCCCGG GGGAGGCAAG
AAAGACTGGT GCCAGTAAGT TTTGTTACTG ATGTCGGGCC CCCTCCGTTC

14001 CACACAGACC ATCAATCTTG ACGACCGGTC GCACTGGGGC GGCGACCTGA
GTGTGTCTGG TAGTTAGAAC TGCTGGCCAG CGTGACCCCG CCGCTGGACT

14051 AAACCATCCT GCATACCAAC ATGCCAAATG TGAACGAGTT CATGTTTACC
TTTGGTAGGA CGTATGGTTG TACGGTTTAC ACTTGCTCAA GTACAAATGG

14101 AATAAGTTTA AGGCGCGGGT GATGGTGTCG CGCTTGCCTA CTAAGGACAA
TTATTCAAAT TCCGCGCCCA CTACCACAGC GCGAACGGAT GATTCCTGTT

14151 TCAGGTGGAG CTGAAATACG AGTGGGTGGA GTTCACGCTG CCCGAGGGCA
AGTCCACCTC GACTTTATGC TCACCCACCT CAAGTGCGAC GGGCTCCCGT

14201 ACTACTCCGA GACCATGACC ATAGACCTTA TGAACAACGC GATCGTGGAG
TGATGAGGCT CTGGTACTGG TATCTGGAAT ACTTGTTGCG CTAGCACCTC

14251 CACTACTTGA AAGTGGGCAG ACAGAACGGG GTTCTGGAAA GCGACATCGG
GTGATGAACT TTCACCCGTC TGTCTTGCCC CAAGACCTTT CGCTGTAGCC

14301 GGTAAAGTTT GACACCCGCA ACTTCAGACT GGGGTTTGAC CCCGTCACTG
CCATTTCAAA CTGTGGGCGT TGAAGTCTGA CCCCAAACTG GGGCAGTGAC

14351 GTCTTGTCAT GCCTGGGGTA TATACAAACG AAGCCTTCCA TCCAGACATC
CAGAACAGTA CGGACCCCAT ATATGTTTGC TTCGGAAGGT AGGTCTGTAG

FIG.9A-17

14401 ATTTTGCTGC CAGGATGCGG GGTGGACTTC ACCCACAGCC GCCTGAGCAA
TAAAACGACG GTCCTACGCC CCACCTGAAG TGGGTGTCGG CGGACTCGTT

14451 CTTGTTGGGC ATCCGCAAGC GGCAACCCTT CCAGGAGGGC TTTAGGATCA
GAACAACCCG TAGGCGTTCG CCGTTGGGAA GGTCCTCCCG AAATCCTAGT

14501 CCTACGATGA TCTGGAGGGT GGTAACATTC CCGCACTGTT GGATGTGGAC
GGATGCTACT AGACCTCCCA CCATTGTAAG GGCGTGACAA CCTACACCTG

14551 GCCTACCAGG CGAGCTTGAA AGATGACACC GAACAGGGCG GGGGTGGCGC
CGGATGGTCC GCTCGAACTT TCTACTGTGG CTTGTCCCGC CCCCACCGCG

14601 AGGCGGCAGC AACAGCAGTG GCAGCGGCGC GGAAGAGAAC TCCAACGCGG
TCCGCCGTCG TTGTCGTCAC CGTCGCCGCG CCTTCTCTTG AGGTTGCGCC

14651 CAGCCGCGGC AATGCAGCCG GTGGAGGACA TGAACGATCA TGCCATTCGC
GTCGGCGCCG TTACGTCGGC CACCTCCTGT ACTTGCTAGT ACGGTAAGCG

14701 GGCGACACCT TTGCCACACG GGCTGAGGAG AAGCGCGCTG AGGCCGAAGC
CCGCTGTGGA AACGGTGTGC CCGACTCCTC TTCGCGCGAC TCCGGCTTCG

14751 AGCGGCCGAA GCTGCCGCCC CCGCTGCGCA ACCCGAGGTC GAGAAGCCTC
TCGCCGGCTT CGACGGCGGG GGCGACGCGT TGGGCTCCAG CTCTTCGGAG

14801 AGAAGAAACC GGTGATCAAA CCCCTGACAG AGGACAGCAA GAAACGCAGT
TCTTCTTTGG CCACTAGTTT GGGGACTGTC TCCTGTCGTT CTTTGCGTCA

14851 TACAACCTAA TAAGCAATGA CAGCACCTTC ACCCAGTACC GCAGCTGGTA
ATGTTGGATT ATTCGTTACT GTCGTGGAAG TGGGTCATGG CGTCGACCAT

14901 CCTTGCATAC AACTACGGCG ACCCTCAGAC CGGAATCCGC TCATGGACCC
GGAACGTATG TTGATGCCGC TGGGAGTCTG GCCTTAGGCG AGTACCTGGG

14951 TGCTTTGCAC TCCTGACGTA ACCTGCGGCT CGGAGCAGGT CTACTGGTCG
ACGAAACGTG AGGACTGCAT TGGACGCCGA GCCTCGTCCA GATGACCAGC

15001 TTGCCAGACA TGATGCAAGA CCCCGTGACC TTCCGCTCCA CGCGCCAGAT
AACGGTCTGT ACTACGTTCT GGGGCACTGG AAGGCGAGGT GCGCGGTCTA

15051 CAGCAACTTT CCGGTGGTGG GCGCCGAGCT GTTGCCCGTG CACTCCAAGA
GTCGTTGAAA GGCCACCACC CGCGGCTCGA CAACGGGCAC GTGAGGTTCT

15101 GCTTCTACAA CGACCAGGCC GTCTACTCCC AACTCATCCG CCAGTTTACC
CGAAGATGTT GCTGGTCCGG CAGATGAGGG TTGAGTAGGC GGTCAAATGG

15151 TCTCTGACCC ACGTGTTCAA TCGCTTTCCC GAGAACCAGA TTTTGGCGCG
AGAGACTGGG TGCACAAGTT AGCGAAAGGG CTCTTGGTCT AAAACCGCGC

15201 CCCGCCAGCC CCCACCATCA CCACCGTCAG TGAAAACGTT CCTGCTCTCA
GGGCGGTCGG GGGTGGTAGT GGTGGCAGTC ACTTTTGCAA GGACGAGAGT

FIG.9A-18

15251 CAGATCACGG GACGCTACCG CTGCGCAACA GCATCGGAGG AGTCCAGCGA
GTCTAGTGCC CTGCGATGGC GACGCGTTGT CGTAGCCTCC TCAGGTCGCT

15301 GTGACCATTA CTGACGCCAG ACGCCGCACC TGCCCCTACG TTTACAAGGC
CACTGGTAAT GACTGCGGTC TGCGGCGTGG ACGGGGATGC AAATGTTCCG

15351 CCTGGGCATA GTCTCGCCGC GCGTCCTATC GAGCCGCACT TTTTGAGCAA
GGACCCGTAT CAGAGCGGCG CGCAGGATAG CTCGGCGTGA AAAACTCGTT

15401 GCATGTCCAT CCTTATATCG CCCAGCAATA ACACAGGCTG GGGCCTGCGC
CGTACAGGTA GGAATATAGC GGGTCGTTAT TGTGTCCGAC CCCGGACGCG

15451 TTCCCAAGCA AGATGTTTGG CGGGGCCAAG AAGCGCTCCG ACCAACACCC
AAGGGTTCGT TCTACAAACC GCCCCGGTTC TTCGCGAGGC TGGTTGTGGG

15501 AGTGCGCGTG CGCGGGCACT ACCGCGCGCC CTGGGGCGCG CACAAACGCG
TCACGCGCAC GCGCCCGTGA TGGCGCGCGG GACCCCGCGC GTGTTTGCGC

15551 GCCGCACTGG GCGCACCACC GTCGATGACG CCATCGACGC GGTGGTGGAG
CGGCGTGACC CGCGTGGTGG CAGCTACTGC GGTAGCTGCG CCACCACCTC

15601 GAGGCGCGCA ACTACACGCC CACGCCGCCA CCAGTGTCCA CAGTGGACGC
CTCCGCGCGT TGATGTGCGG GTGCGGCGGT GGTCACAGGT GTCACCTGCG

15651 GGCCATTCAG ACCGTGGTGC GCGGAGCCCG GCGCTATGCT AAAATGAAGA
CCGGTAAGTC TGGCACCACG CGCCTCGGGC CGCGATACGA TTTTACTTCT

15701 GACGGCGGAG GCGCGTAGCA CGTCGCCACC GCCGCCGACC CGGCACTGCC
CTGCCGCCTC CGCGCATCGT GCAGCGGTGG CGGCGGCTGG GCCGTGACGG

15751 GCCCAACGCG CGGCGGCGGC CCTGCTTAAC CGCGCACGTC GCACCGGCCG
CGGGTTGCGC GCCGCCGCCG GGACGAATTG GCGCGTGCAG CGTGGCCGGC

15801 ACGGGCGGCC ATGCGGGCCG CTCGAAGGCT GGCCGCGGGT ATTGTCACTG
TGCCCGCCGG TACGCCCGGC GAGCTTCCGA CCGGCGCCCA TAACAGTGAC

15851 TGCCCCCCAG GTCCAGGCGA CGAGCGGCCG CCGCAGCAGC CGCGGCCATT
ACGGGGGGTC CAGGTCCGCT GCTCGCCGGC GGCGTCGTCG GCGCCGGTAA

15901 AGTGCTATGA CTCAGGGTCG CAGGGGCAAC GTGTATTGGG TGCGCGACTC
TCACGATACT GAGTCCCAGC GTCCCCGTTG CACATAACCC ACGCGCTGAG

15951 GGTTAGCGGC CTGCGCGTGC CCGTGCGCAC CCGCCCCCCG CGCAACTAGA
CCAATCGCCG GACGCGCACG GGCACGCGTG GGCGGGGGGC GCGTTGATCT

16001 TTGCAAGAAA AAACTACTTA GACTCGTACT GTTGTATGTA TCCAGCGGCG
AACGTTCTTT TTTGATGAAT CTGAGCATGA CAACATACAT AGGTCGCCGC

16051 GCGGCGCGCA ACGAAGCTAT GTCCAAGCGC AAAATCAAAG AAGAGATGCT
CGCCGCGCGT TGCTTCGATA CAGGTTCGCG TTTTAGTTTC TTCTCTACGA

FIG.9A-19

16101 CCAGGTCATC GCGCCGGAGA TCTATGGCCC CCCGAAGAAG GAAGAGCAGG
GGTCCAGTAG CGCGGCCTCT AGATACCGGG GGGCTTCTTC CTTCTCGTCC

16151 ATTACAAGCC CCGAAAGCTA AAGCGGGTCA AAAAGAAAAA GAAAGATGAT
TAATGTTCGG GGCTTTCGAT TTCGCCCAGT TTTTCTTTTT CTTTCTACTA

16201 GATGATGAAC TTGACGACGA GGTGGAACTG CTGCACGCTA CCGCGCCCAG
CTACTACTTG AACTGCTGCT CCACCTTGAC GACGTGCGAT GGCGCGGGTC

16251 GCGACGGGTA CAGTGGAAAG GTCGACGCGT AAAACGTGTT TTGCGACCCG
CGCTGCCCAT GTCACCTTTC CAGCTGCGCA TTTTGCACAA AACGCTGGGC

16301 GCACCACCGT AGTCTTTACG CCCGGTGAGC GCTCCACCCG CACCTACAAG
CGTGGTGGCA TCAGAAATGC GGGCCACTCG CGAGGTGGGC GTGGATGTTC

16351 CGCGTGTATG ATGAGGTGTA CGGCGACGAG GACCTGCTTG AGCAGGCCAA
GCGCACATAC TACTCCACAT GCCGCTGCTC CTGGACGAAC TCGTCCGGTT

16401 CGAGCGCCTC GGGGAGTTTG CCTACGGAAA GCGGCATAAG GACATGCTGG
GCTCGCGGAG CCCCTCAAAC GGATGCCTTT CGCCGTATTC CTGTACGACC

16451 CGTTGCCGCT GGACGAGGGC AACCCAACAC CTAGCCTAAA GCCCGTAACA
GCAACGGCGA CCTGCTCCCG TTGGGTTGTG GATCGGATTT CGGGCATTGT

16501 CTGCAGCAGG TGCTGCCCGC GCTTGCACCG TCCGAAGAAA AGCGCGGCCT
GACGTCGTCC ACGACGGGCG CGAACGTGGC AGGCTTCTTT TCGCGCCGGA

16551 AAAGCGCGAG TCTGGTGACT TGGCACCCAC CGTGCAGCTG ATGGTACCCA
TTTCGCGCTC AGACCACTGA ACCGTGGGTG GCACGTCGAC TACCATGGGT

16601 AGCGCCAGCG ACTGGAAGAT GTCTTGGAAA AAATGACCGT GGAACCTGGG
TCGCGGTCGC TGACCTTCTA CAGAACCTTT TTTACTGGCA CCTTGGACCC

16651 CTGGAGCCCG AGGTCCGCGT GCGGCCAATC AAGCAGGTGG CGCCGGGACT
GACCTCGGGC TCCAGGCGCA CGCCGGTTAG TTCGTCCACC GCGGCCCTGA

16701 GGGCGTGCAG ACCGTGGACG TTCAGATACC CACTACCAGT AGCACCAGTA
CCCGCACGTC TGGCACCTGC AAGTCTATGG GTGATGGTCA TCGTGGTCAT

16751 TTGCCACCGC CACAGAGGGC ATGGAGACAC AAACGTCCCC GGTTGCCTCA
AACGGTGGCG GTGTCTCCCG TACCTCTGTG TTTGCAGGGG CCAACGGAGT

16801 GCGGTGGCGG ATGCCGCGGT GCAGGCGGTC GCTGCGGCCG CGTCCAAGAC
CGCCACCGCC TACGGCGCCA CGTCCGCCAG CGACGCCGGC GCAGGTTCTG

16851 CTCTACGGAG GTGCAAACGG ACCCGTGGAT GTTTCGCGTT TCAGCCCCCC
GAGATGCCTC CACGTTTGCC TGGGCACCTA CAAAGCGCAA AGTCGGGGGG

16901 GGCGCCCGCG CCGTTCGAGG AAGTACGGCG CCGCCAGCGC GCTACTGCCC
CCGCGGGCGC GGCAAGCTCC TTCATGCCGC GGCGGTCGCG CGATGACGGG

FIG.9A-20

16951 GAATATGCCC TACATCCTTC CATTGCGCCT ACCCCCGGCT ATCGTGGCTA
CTTATACGGG ATGTAGGAAG GTAACGCGGA TGGGGGCCGA TAGCACCGAT

17001 CACCTACCGC CCCAGAAGAC GAGCAACTAC CCGACGCCGA ACCACCACTG
GTGGATGGCG GGGTCTTCTG CTCGTTGATG GGCTGCGGCT TGGTGGTGAC

17051 GAACCCGCCG CCGCCGTCGC CGTCGCCAGC CCGTGCTGGC CCCGATTTCC
CTTGGGCGGC GGCGGCAGCG GCAGCGGTCG GGCACGACCG GGGCTAAAGG

17101 GTGCGCAGGG TGGCTCGCGA AGGAGGCAGG ACCCTGGTGC TGCCAACAGC
CACGCGTCCC ACCGAGCGCT TCCTCCGTCC TGGGACCACG ACGGTTGTCG

17151 GCGCTACCAC CCCAGCATCG TTTAAAAGCC GGTCTTTGTG GTTCTTGCAG
CGCGATGGTG GGGTCGTAGC AAATTTTCGG CCAGAAACAC CAAGAACGTC

17201 ATATGGCCCT CACCTGCCGC CTCCGTTTCC CGGTGCCGGG ATTCCGAGGA
TATACCGGGA GTGGACGGCG GAGGCAAAGG GCCACGGCCC TAAGGCTCCT

17251 AGAATGCACC GTAGGAGGGG CATGGCCGGC CACGGCCTGA CGGGCGGCAT
TCTTACGTGG CATCCTCCCC GTACCGGCCG GTGCCGGACT GCCCGCCGTA

17301 GCGTCGTGCG CACCACCGGC GGCGGCGCGC GTCGCACCGT CGCATGCGCG
CGCAGCACGC GTGGTGGCCG CCGCCGCGCG CAGCGTGGCA GCGTACGCGC

17351 GCGGTATCCT GCCCCTCCTT ATTCCACTGA TCGCCGCGGC GATTGGCGCC
CGCCATAGGA CGGGGAGGAA TAAGGTGACT AGCGGCGCCG CTAACCGCGG

17401 GTGCCCGGAA TTGCATCCGT GGCCTTGCAG GCGCAGAGAC ACTGATTAAA
CACGGGCCTT AACGTAGGCA CCGGAACGTC CGCGTCTCTG TGACTAATTT

17451 AACAAGTTGC ATGTGGAAAA ATCAAAATAA AAAGTCTGGA CTCTCACGCT
TTGTTCAACG TACACCTTTT TAGTTTTATT TTTCAGACCT GAGAGTGCGA

17501 CGCTTGGTCC TGTAACTATT TTGTAGAATG GAAGACATCA ACTTTGCGTC
GCGAACCAGG ACATTGATAA AACATCTTAC CTTCTGTAGT TGAAACGCAG

17551 TCTGGCCCCG CGACACGGCT CGCGCCCGTT CATGGGAAAC TGGCAAGATA
AGACCGGGGC GCTGTGCCGA GCGCGGGCAA GTACCCTTTG ACCGTTCTAT

17601 TCGGCACCAG CAATATGAGC GGTGGCGCCT TCAGCTGGGG CTCGCTGTGG
AGCCGTGGTC GTTATACTCG CCACCGCGGA AGTCGACCCC GAGCGACACC

17651 AGCGGCATTA AAAATTTCGG TTCCACCGTT AAGAACTATG GCAGCAAGGC
TCGCCGTAAT TTTTAAAGCC AAGGTGGCAA TTCTTGATAC CGTCGTTCCG

17701 CTGGAACAGC AGCACAGGCC AGATGCTGAG GGATAAGTTG AAAGAGCAAA
GACCTTGTCG TCGTGTCCGG TCTACGACTC CCTATTCAAC TTTCTCGTTT

17751 ATTTCCAACA AAAGGTGGTA GATGGCCTGG CCTCTGGCAT TAGCGGGGTG
TAAAGGTTGT TTTCCACCAT CTACCGGACC GGAGACCGTA ATCGCCCCAC

FIG.9A-21

17801 GTGGACCTGG CCAACCAGGC AGTGCAAAAT AAGATTAACA GTAAGCTTGA
CACCTGGACC GGTTGGTCCG TCACGTTTTA TTCTAATTGT CATTCGAACT

17851 TCCCCGCCCT CCCGTAGAGG AGCCTCCACC GGCCGTGGAG ACAGTGTCTC
AGGGGCGGGA GGGCATCTCC TCGGAGGTGG CCGGCACCTC TGTCACAGAG

17901 CAGAGGGGCG TGGCGAAAAG CGTCCGCGCC CCGACAGGGA AGAAACTCTG
GTCTCCCCGC ACCGCTTTTC GCAGGCGCGG GGCTGTCCCT TCTTTGAGAC

17951 GTGACGCAAA TAGACGAGCC TCCCTCGTAC GAGGAGGCAC TAAAGCAAGG
CACTGCGTTT ATCTGCTCGG AGGGAGCATG CTCCTCCGTG ATTTCGTTCC

18001 CCTGCCCACC ACCCGTCCCA TCGCGCCCAT GGCTACCGGA GTGCTGGGCC
GGACGGGTGG TGGGCAGGGT AGCGCGGGTA CCGATGGCCT CACGACCCGG

18051 AGCACACACC CGTAACGCTG GACCTGCCTC CCCCCGCCGA CACCCAGCAG
TCGTGTGTGG GCATTGCGAC CTGGACGGAG GGGGGCGGCT GTGGGTCGTC

18101 AAACCTGTGC TGCCAGGCCC GACCGCCGTT GTTGTAACCC GTCCTAGCCG
TTTGGACACG ACGGTCCGGG CTGGCGGCAA CAACATTGGG CAGGATCGGC

18151 CGCGTCCCTG CGCCGCGCCG CCAGCGGTCC GCGATCGTTG CGGCCCGTAG
GCGCAGGGAC GCGGCGCGGC GGTCGCCAGG CGCTAGCAAC GCCGGGCATC

18201 CCAGTGGCAA CTGGCAAAGC ACACTGAACA GCATCGTGGG TCTGGGGGTG
GGTCACCGTT GACCGTTTCG TGTGACTTGT CGTAGCACCC AGACCCCCAC

18251 CAATCCCTGA AGCGCCGACG ATGCTTCTGA TAGCTAACGT GTCGTATGTG
GTTAGGGACT TCGCGGCTGC TACGAAGACT ATCGATTGCA CAGCATACAC

18301 TGTCATGTAT GCGTCCATGT CGCCGCCAGA GGAGCTGCTG AGCCGCCGCG
ACAGTACATA CGCAGGTACA GCGGCGGTCT CCTCGACGAC TCGGCGGCGC

18351 CGCCCGCTTT CCAAGATGGC TACCCCTTCG ATGATGCCGC AGTGGTCTTA
GCGGGCGAAA GGTTCTACCG ATGGGGAAGC TACTACGGCG TCACCAGAAT

18401 CATGCACATC TCGGGCCAGG ACGCCTCGGA GTACCTGAGC CCCGGGCTGG
GTACGTGTAG AGCCCGGTCC TGCGGAGCCT CATGGACTCG GGGCCCGACC

18451 TGCAGTTTGC CCGCGCCACC GAGACGTACT TCAGCCTGAA TAACAAGTTT
ACGTCAAACG GGCGCGGTGG CTCTGCATGA AGTCGGACTT ATTGTTCAAA

18501 AGAAACCCCA CGGTGGCGCC TACGCACGAC GTGACCACAG ACCGGTCCCA
TCTTTGGGGT GCCACCGCGG ATGCGTGCTG CACTGGTGTC TGGCCAGGGT

18551 GCGTTTGACG CTGCGGTTCA TCCCTGTGGA CCGTGAGGAT ACTGCGTACT
CGCAAACTGC GACGCCAAGT AGGGACACCT GGCACTCCTA TGACGCATGA

18601 CGTACAAGGC GCGGTTCACC CTAGCTGTGG GTGATAACCG TGTGCTGGAC
GCATGTTCCG CGCCAAGTGG GATCGACACC CACTATTGGC ACACGACCTG

FIG.9A-22

18651 ATGGCTTCCA CGTACTTTGA CATCCGCGGC GTGCTGGACA GGGGCCCTAC
TACCGAAGGT GCATGAAACT GTAGGCGCCG CACGACCTGT CCCCGGGATG

18701 TTTTAAGCCC TACTCTGGCA CTGCCTACAA CGCCCTGGCT CCCAAGGGTG
AAAATTCGGG ATGAGACCGT GACGGATGTT GCGGGACCGA GGGTTCCCAC

18751 CCCCAAATCC TTGCGAATGG GATGAAGCTG CTACTGCTCT TGAAATAAAC
GGGGTTTAGG AACGCTTACC CTACTTCGAC GATGACGAGA ACTTTATTTG

18801 CTAGAAGAAG AGGACGATGA CAACGAAGAC GAAGTAGACG AGCAAGCTGA
GATCTTCTTC TCCTGCTACT GTTGCTTCTG CTTCATCTGC TCGTTCGACT

18851 GCAGCAAAAA ACTCACGTAT TTGGGCAGGC GCCTTATTCT GGTATAAATA
CGTCGTTTTT TGAGTGCATA AACCCGTCCG CGGAATAAGA CCATATTTAT

18901 TTACAAAGGA GGGTATTCAA ATAGGTGTCG AAGGTCAAAC ACCTAAATAT
AATGTTTCCT CCCATAAGTT TATCCACAGC TTCCAGTTTG TGGATTTATA

18951 GCCGATAAAA CATTTCAACC TGAACCTCAA ATAGGAGAAT CTCAGTGGTA
CGGCTATTTT GTAAAGTTGG ACTTGGAGTT TATCCTCTTA GAGTCACCAT

19001 CGAAACAGAA ATTAATCATG CAGCTGGGAG AGTCCTAAAA AAGACTACCC
GCTTTGTCTT TAATTAGTAC GTCGACCCTC TCAGGATTTT TTCTGATGGG

19051 CAATGAAACC ATGTTACGGT TCATATGCAA AACCCACAAA TGAAAATGGA
GTTACTTTGG TACAATGCCA AGTATACGTT TTGGGTGTTT ACTTTTACCT

19101 GGGCAAGGCA TTCTTGTAAA GCAACAAAAT GGAAAGCTAG AAAGTCAAGT
CCCGTTCCGT AAGAACATTT CGTTGTTTTA CCTTTCGATC TTTCAGTTCA

19151 GGAAATGCAA TTTTTCTCAA CTACTGAGGC AGCCGCAGGC AATGGTGATA
CCTTTACGTT AAAAAGAGTT GATGACTCCG TCGGCGTCCG TTACCACTAT

19201 ACTTGACTCC TAAAGTGGTA TTGTACAGTG AAGATGTAGA TATAGAAACC
TGAACTGAGG ATTTCACCAT AACATGTCAC TTCTACATCT ATATCTTTGG

19251 CCAGACACTC ATATTTCTTA CATGCCCACT ATTAAGGAAG GTAACTCACG
GGTCTGTGAG TATAAAGAAT GTACGGGTGA TAATTCCTTC CATTGAGTGC

19301 AGAACTAATG GGCCAACAAT CTATGCCCAA CAGGCCTAAT TACATTGCTT
TCTTGATTAC CCGGTTGTTA GATACGGGTT GTCCGGATTA ATGTAACGAA

19351 TTAGGGACAA TTTTATTGGT CTAATGTATT ACAACAGCAC GGGTAATATG
AATCCCTGTT AAAATAACCA GATTACATAA TGTTGTCGTG CCCATTATAC

19401 GGTGTTCTGG CGGGCCAAGC ATCGCAGTTG AATGCTGTTG TAGATTTGCA
CCACAAGACC GCCCGGTTCG TAGCGTCAAC TTACGACAAC ATCTAAACGT

19451 AGACAGAAAC ACAGAGCTTT CATACCAGCT TTTGCTTGAT TCCATTGGTG
TCTGTCTTTG TGTCTCGAAA GTATGGTCGA AAACGAACTA AGGTAACCAC

FIG.9A-23

19501 ATAGAACCAG GTACTTTTCT ATGTGGAATC AGGCTGTTGA CAGCTATGAT
TATCTTGGTC CATGAAAAGA TACACCTTAG TCCGACAACT GTCGATACTA

19551 CCAGATGTTA GAATTATTGA AAATCATGGA ACTGAAGATG AACTTCCAAA
GGTCTACAAT CTTAATAACT TTTAGTACCT TGACTTCTAC TTGAAGGTTT

19601 TTACTGCTTT CCACTGGGAG GTGTGATTAA TACAGAGACT CTTACCAAGG
AATGACGAAA GGTGACCCTC CACACTAATT ATGTCTCTGA GAATGGTTCC

19651 TAAAACCTAA AACAGGTCAG GAAAATGGAT GGGAAAAAGA TGCTACAGAA
ATTTTGGATT TTGTCCAGTC CTTTTACCTA CCCTTTTTCT ACGATGTCTT

19701 TTTTCAGATA AAAATGAAAT AAGAGTTGGA AATAATTTTG CCATGGAAAT
AAAAGTCTAT TTTTACTTTA TTCTCAACCT TTATTAAAAC GGTACCTTTA

19751 CAATCTAAAT GCCAACCTGT GGAGAAATTT CCTGTACTCC AACATAGCGC
GTTAGATTTA CGGTTGGACA CCTCTTTAAA GGACATGAGG TTGTATCGCG

19801 TGTATTTGCC CGACAAGCTA AAGTACAGTC CTTCCAACGT AAAAATTTCT
ACATAAACGG GCTGTTCGAT TTCATGTCAG GAAGGTTGCA TTTTTAAAGA

19851 GATAACCCAA ACACCTACGA CTACATGAAC AAGCGAGTGG TGGCTCCCGG
CTATTGGGTT TGTGGATGCT GATGTACTTG TTCGCTCACC ACCGAGGGCC

19901 GCTAGTGGAC TGCTACATTA ACCTTGGAGC ACGCTGGTCC CTTGACTATA
CGATCACCTG ACGATGTAAT TGGAACCTCG TGCGACCAGG GAACTGATAT

19951 TGGACAACGT CAACCCATTT AACCACCACC GCAATGCTGG CCTGCGCTAC
ACCTGTTGCA GTTGGGTAAA TTGGTGGTGG CGTTACGACC GGACGCGATG

20001 CGCTCAATGT TGCTGGGCAA TGGTCGCTAT GTGCCCTTCC ACATCCAGGT
GCGAGTTACA ACGACCCGTT ACCAGCGATA CACGGGAAGG TGTAGGTCCA

20051 GCCTCAGAAG TTCTTTGCCA TTAAAAACCT CCTTCTCCTG CCGGGCTCAT
CGGAGTCTTC AAGAAACGGT AATTTTTGGA GGAAGAGGAC GGCCCGAGTA

20101 ACACCTACGA GTGGAACTTC AGGAAGGATG TTAACATGGT TCTGCAGAGC
TGTGGATGCT CACCTTGAAG TCCTTCCTAC AATTGTACCA AGACGTCTCG

20151 TCCCTAGGAA ATGACCTAAG GGTTGACGGA GCCAGCATTA AGTTTGATAG
AGGGATCCTT TACTGGATTC CCAACTGCCT CGGTCGTAAT TCAAACTATC

20201 CATTTGCCTT TACGCCACCT TCTTCCCCAT GGCCCACAAC ACCGCCTCCA
GTAAACGGAA ATGCGGTGGA AGAAGGGGTA CCGGGTGTTG TGGCGGAGGT

20251 CGCTTGAGGC CATGCTTAGA AACGACACCA ACGACCAGTC CTTTAACGAC
GCGAACTCCG GTACGAATCT TTGCTGTGGT TGCTGGTCAG GAAATTGCTG

20301 TATCTCTCCG CCGCCAACAT GCTCTACCCT ATACCCGCCA ACGCTACCAA
ATAGAGAGGC GGCGGTTGTA CGAGATGGGA TATGGGCGGT TGCGATGGTT

FIG.9A-24

20351 CGTGCCCATA TCCATCCCCT CCCGCAACTG GGCGGCTTTC CGCGGCTGGG
GCACGGGTAT AGGTAGGGGA GGGCGTTGAC CCGCCGAAAG GCGCCGACCC

20401 CCTTCACGCG CCTTAAGACT AAGGAAACCC CATCACTGGG CTCGGGCTAC
GGAAGTGCGC GGAATTCTGA TTCCTTTGGG GTAGTGACCC GAGCCCGATG

20451 GACCCTTATT ACACCTACTC TGGCTCTATA CCCTACCTAG ATGGAACCTT
CTGGGAATAA TGTGGATGAG ACCGAGATAT GGGATGGATC TACCTTGGAA

20501 TTACCTCAAC CACACCTTTA AGAAGGTGGC CATTACCTTT GACTCTTCTG
AATGGAGTTG GTGTGGAAAT TCTTCCACCG GTAATGGAAA CTGAGAAGAC

20551 TCAGCTGGCC TGGCAATGAC CGCCTGCTTA CCCCCAACGA GTTTGAAATT
AGTCGACCGG ACCGTTACTG GCGGACGAAT GGGGGTTGCT CAAACTTTAA

20601 AAGCGCTCAG TTGACGGGGA GGGTTACAAC GTTGCCCAGT GTAACATGAC
TTCGCGAGTC AACTGCCCCT CCCAATGTTG CAACGGGTCA CATTGTACTG

20651 CAAAGACTGG TTCCTGGTAC AAATGCTAGC TAACTATAAC ATTGGCTACC
GTTTCTGACC AAGGACCATG TTTACGATCG ATTGATATTG TAACCGATGG

20701 AGGGCTTCTA TATCCCAGAG AGCTACAAGG ACCGCATGTA CTCCTTCTTT
TCCCGAAGAT ATAGGGTCTC TCGATGTTCC TGGCGTACAT GAGGAAGAAA

20751 AGAAACTTCC AGCCCATGAG CCGTCAGGTG GTGGATGATA CTAAATACAA
TCTTTGAAGG TCGGGTACTC GGCAGTCCAC CACCTACTAT GATTTATGTT

20801 GGACTACCAA CAGGTGGGCA TCCTACACCA ACACAACAAC TCTGGATTTG
CCTGATGGTT GTCCACCCGT AGGATGTGGT TGTGTTGTTG AGACCTAAAC

20851 TTGGCTACCT TGCCCCCACC ATGCGCGAAG GACAGGCCTA CCCTGCTAAC
AACCGATGGA ACGGGGGTGG TACGCGCTTC CTGTCCGGAT GGGACGATTG

20901 TTCCCCTATC CGCTTATAGG CAAGACCGCA GTTGACAGCA TTACCCAGAA
AAGGGGATAG GCGAATATCC GTTCTGGCGT CAACTGTCGT AATGGGTCTT

20951 AAAGTTTCTT TGCGATCGCA CCCTTTGGCG CATCCCATTC TCCAGTAACT
TTTCAAAGAA ACGCTAGCGT GGGAAACCGC GTAGGGTAAG AGGTCATTGA

21001 TTATGTCCAT GGGCGCACTC ACAGACCTGG GCCAAAACCT TCTCTACGCC
AATACAGGTA CCCGCGTGAG TGTCTGGACC CGGTTTTGGA AGAGATGCGG

21051 AACTCCGCCC ACGCGCTAGA CATGACTTTT GAGGTGGATC CCATGGACGA
TTGAGGCGGG TGCGCGATCT GTACTGAAAA CTCCACCTAG GGTACCTGCT

21101 GCCCACCCTT CTTTATGTTT TGTTTGAAGT CTTTGACGTG GTCCGTGTGC
CGGGTGGGAA GAAATACAAA ACAAACTTCA GAAACTGCAC CAGGCACACG

21151 ACCAGCCGCA CCGCGGCGTC ATCGAAACCG TGTACCTGCG CACGCCCTTC
TGGTCGGCGT GGCGCCGCAG TAGCTTTGGC ACATGGACGC GTGCGGGAAG

FIG.9A-25

21201 TCGGCCGGCA ACGCCACAAC ATAAAGAAGC AAGCAACATC AACAACAGCT
AGCCGGCCGT TGCGGTGTTG TATTTCTTCG TTCGTTGTAG TTGTTGTCGA

21251 GCCGCCATGG GCTCCAGTGA GCAGGAACTG AAAGCCATTG TCAAAGATCT
CGGCGGTACC CGAGGTCACT CGTCCTTGAC TTTCGGTAAC AGTTTCTAGA

21301 TGGTTGTGGG CCATATTTTT TGGGCACCTA TGACAAGCGC TTTCCAGGCT
ACCAACACCC GGTATAAAAA ACCCGTGGAT ACTGTTCGCG AAAGGTCCGA

21351 TTGTTTCTCC ACACAAGCTC GCCTGCGCCA TAGTCAATAC GGCCGGTCGC
AACAAAGAGG TGTGTTCGAG CGGACGCGGT ATCAGTTATG CCGGCCAGCG

21401 GAGACTGGGG GCGTACACTG GATGGCCTTT GCCTGGAACC CGCACTCAAA
CTCTGACCCC CGCATGTGAC CTACCGGAAA CGGACCTTGG GCGTGAGTTT

21451 AACATGCTAC CTCTTTGAGC CCTTTGGCTT TTCTGACCAG CGACTCAAGC
TTGTACGATG GAGAAACTCG GGAAACCGAA AAGACTGGTC GCTGAGTTCG

21501 AGGTTTACCA GTTTGAGTAC GAGTCACTCC TGCGCCGTAG CGCCATTGCT
TCCAAATGGT CAAACTCATG CTCAGTGAGG ACGCGGCATC GCGGTAACGA

21551 TCTTCCCCCG ACCGCTGTAT AACGCTGGAA AAGTCCACCC AAAGCGTACA
AGAAGGGGGC TGGCGACATA TTGCGACCTT TTCAGGTGGG TTTCGCATGT

21601 GGGGCCCAAC TCGGCCGCCT GTGGACTATT CTGCTGCATG TTTCTCCACG
CCCCGGGTTG AGCCGGCGGA CACCTGATAA GACGACGTAC AAAGAGGTGC

21651 CCTTTGCCAA CTGGCCCCAA ACTCCCATGG ATCACAACCC CACCATGAAC
GGAAACGGTT GACCGGGGTT TGAGGGTACC TAGTGTTGGG GTGGTACTTG

21701 CTTATTACCG GGGTACCCAA CTCCATGCTC AACAGTCCCC AGGTACAGCC
GAATAATGGC CCCATGGGTT GAGGTACGAG TTGTCAGGGG TCCATGTCGG

21751 CACCCTGCGT CGCAACCAGG AACAGCTCTA CAGCTTCCTG GAGCGCCACT
GTGGGACGCA GCGTTGGTCC TTGTCGAGAT GTCGAAGGAC CTCGCGGTGA

21801 CGCCCTACTT CCGCAGCCAC AGTGCGCAGA TTAGGAGCGC CACTTCTTTT
GCGGGATGAA GGCGTCGGTG TCACGCGTCT AATCCTCGCG GTGAAGAAAA

21851 TGTCACTTGA AAAACATGTA AAAATAATGT ACTAGAGACA CTTTCAATAA
ACAGTGAACT TTTTGTACAT TTTTATTACA TGATCTCTGT GAAAGTTATT

21901 AGGCAAATGC TTTTATTTGT ACACTCTCGG GTGATTATTT ACCCCCACCC
TCCGTTTACG AAAATAAACA TGTGAGAGCC CACTAATAAA TGGGGGTGGG

21951 TTGCCGTCTG CGCCGTTTAA AAATCAAAGG GGTTCTGCCG CGCATCGCTA
AACGGCAGAC GCGGCAAATT TTTAGTTTCC CCAAGACGGC GCGTAGCGAT

22001 TGCGCCACTG GCAGGGACAC GTTGCGATAC TGGTGTTTAG TGCTCCACTT
ACGCGGTGAC CGTCCCTGTG CAACGCTATG ACCACAAATC ACGAGGTGAA

FIG.9A-26

22051 AAACTCAGGC ACAACCATCC GCGGCAGCTC GGTGAAGTTT TCACTCCACA
TTTGAGTCCG TGTTGGTAGG CGCCGTCGAG CCACTTCAAA AGTGAGGTGT

22101 GGCTGCGCAC CATCACCAAC GCGTTTAGCA GGTCGGGCGC CGATATCTTG
CCGACGCGTG GTAGTGGTTG CGCAAATCGT CCAGCCCGCG GCTATAGAAC

22151 AAGTCGCAGT TGGGGCCTCC GCCCTGCGCG CGCGAGTTGC GATACACAGG
TTCAGCGTCA ACCCCGGAGG CGGGACGCGC GCGCTCAACG CTATGTGTCC

22201 GTTGCAGCAC TGGAACACTA TCAGCGCCGG GTGGTGCACG CTGGCCAGCA
CAACGTCGTG ACCTTGTGAT AGTCGCGGCC CACCACGTGC GACCGGTCGT

22251 CGCTCTTGTC GGAGATCAGA TCCGCGTCCA GGTCCTCCGC GTTGCTCAGG
GCGAGAACAG CCTCTAGTCT AGGCGCAGGT CCAGGAGGCG CAACGAGTCC

22301 GCGAACGGAG TCAACTTTGG TAGCTGCCTT CCCAAAAAGG GCGCGTGCCC
CGCTTGCCTC AGTTGAAACC ATCGACGGAA GGGTTTTTCC CGCGCACGGG

22351 AGGCTTTGAG TTGCACTCGC ACCGTAGTGG CATCAAAAGG TGACCGTGCC
TCCGAAACTC AACGTGAGCG TGGCATCACC GTAGTTTTCC ACTGGCACGG

22401 CGGTCTGGGC GTTAGGATAC AGCGCCTGCA TAAAAGCCTT GATCTGCTTA
GCCAGACCCG CAATCCTATG TCGCGGACGT ATTTTCGGAA CTAGACGAAT

22451 AAAGCCACCT GAGCCTTTGC GCCTTCAGAG AAGAACATGC CGCAAGACTT
TTTCGGTGGA CTCGGAAACG CGGAAGTCTC TTCTTGTACG GCGTTCTGAA

22501 GCCGGAAAAC TGATTGGCCG GACAGGCCGC GTCGTGCACG CAGCACCTTG
CGGCCTTTTG ACTAACCGGC CTGTCCGGCG CAGCACGTGC GTCGTGGAAC

22551 CGTCGGTGTT GGAGATCTGC ACCACATTTC GGCCCCACCG GTTCTTCACG
GCAGCCACAA CCTCTAGACG TGGTGTAAAG CCGGGGTGGC CAAGAAGTGC

22601 ATCTTGGCCT TGCTAGACTG CTCCTTCAGC GCGCGCTGCC CGTTTTCGCT
TAGAACCGGA ACGATCTGAC GAGGAAGTCG CGCGCGACGG GCAAAAGCGA

22651 CGTCACATCC ATTTCAATCA CGTGCTCCTT ATTTATCATA ATGCTTCCGT
GCAGTGTAGG TAAAGTTAGT GCACGAGGAA TAAATAGTAT TACGAAGGCA

22701 GTAGACACTT AAGCTCGCCT TCGATCTCAG CGCAGCGGTG CAGCCACAAC
CATCTGTGAA TTCGAGCGGA AGCTAGAGTC GCGTCGCCAC GTCGGTGTTG

22751 GCGCAGCCCG TGGGCTCGTG ATGCTTGTAG GTCACCTCTG CAAACGACTG
CGCGTCGGGC ACCCGAGCAC TACGAACATC CAGTGGAGAC GTTTGCTGAC

22801 CAGGTACGCC TGCAGGAATC GCCCCATCAT CGTCACAAAG GTCTTGTTGC
GTCCATGCGG ACGTCCTTAG CGGGGTAGTA GCAGTGTTTC CAGAACAACG

22851 TGGTGAAGGT CAGCTGCAAC CCGCGGTGCT CCTCGTTCAG CCAGGTCTTG
ACCACTTCCA GTCGACGTTG GGCGCCACGA GGAGCAAGTC GGTCCAGAAC

FIG.9A-27

22901 CATACGGCCG CCAGAGCTTC CACTTGGTCA GGCAGTAGTT TGAAGTTCGC
GTATGCCGGC GGTCTCGAAG GTGAACCAGT CCGTCATCAA ACTTCAAGCG

22951 CTTTAGATCG TTATCCACGT GGTACTTGTC CATCAGCGCG CGCGCAGCCT
GAAATCTAGC AATAGGTGCA CCATGAACAG GTAGTCGCGC GCGCGTCGGA

23001 CCATGCCCTT CTCCCACGCA GACACGATCG GCACACTCAG CGGGTTCATC
GGTACGGGAA GAGGGTGCGT CTGTGCTAGC CGTGTGAGTC GCCCAAGTAG

23051 ACCGTAATTT CACTTTCCGC TTCGCTGGGC TCTTCCTCTT CCTCTTGCGT
TGGCATTAAA GTGAAAGGCG AAGCGACCCG AGAAGGAGAA GGAGAACGCA

23101 CCGCATACCA CGCGCCACTG GGTCGTCTTC ATTCAGCCGC CGCACTGTGC
GGCGTATGGT GCGCGGTGAC CCAGCAGAAG TAAGTCGGCG GCGTGACACG

23151 GCTTACCTCC TTTGCCATGC TTGATTAGCA CCGGTGGGTT GCTGAAACCC
CGAATGGAGG AAACGGTACG AACTAATCGT GGCCACCCAA CGACTTTGGG

23201 ACCATTTGTA GCGCCACATC TTCTCTTTCT TCCTCGCTGT CCACGATTAC
TGGTAAACAT CGCGGTGTAG AAGAGAAAGA AGGAGCGACA GGTGCTAATG

23251 CTCTGGTGAT GGCGGGCGCT CGGGCTTGGG AGAAGGGCGC TTCTTTTTCT
GAGACCACTA CCGCCCGCGA GCCCGAACCC TCTTCCCGCG AAGAAAAAGA

23301 TCTTGGGCGC AATGGCCAAA TCCGCCGCCG AGGTCGATGG CCGCGGGCTG
AGAACCCGCG TTACCGGTTT AGGCGGCGGC TCCAGCTACC GGCGCCCGAC

23351 GGTGTGCGCG GCACCAGCGC GTCTTGTGAT GAGTCTTCCT CGTCCTCGGA
CCACACGCGC CGTGGTCGCG CAGAACACTA CTCAGAAGGA GCAGGAGCCT

23401 CTCGATACGC CGCCTCATCC GCTTTTTTGG GGGCGCCCGG GGAGGCGGCG
GAGCTATGCG GCGGAGTAGG CGAAAAAACC CCCGCGGGCC CCTCCGCCGC

23451 GCGACGGGGA CGGGGACGAC ACGTCCTCCA TGGTTGGGGG ACGTCGCGCC
CGCTGCCCCT GCCCCTGCTG TGCAGGAGGT ACCAACCCCC TGCAGCGCGG

23501 GCACCGCGTC CGCGCTCGGG GGTGGTTTCG CGCTGCTCCT CTTCCCGACT
CGTGGCGCAG GCGCGAGCCC CCACCAAAGC GCGACGAGGA GAAGGGCTGA

23551 GGCCATTTCC TTCTCCTATA GGCAGAAAAA GATCATGGAG TCAGTCGAGA
CCGGTAAAGG AAGAGGATAT CCGTCTTTTT CTAGTACCTC AGTCAGCTCT

23601 AGAAGGACAG CCTAACCGCC CCCTCTGAGT TCGCCACCAC CGCCTCCACC
TCTTCCTGTC GGATTGGCGG GGGAGACTCA AGCGGTGGTG GCGGAGGTGG

23651 GATGCCGCCA ACGCGCCTAC CACCTTCCCC GTCGAGGCAC CCCCGCTTGA
CTACGGCGGT TGCGCGGATG GTGGAAGGGG CAGCTCCGTG GGGGCGAACT

23701 GGAGGAGGAA GTGATTATCG AGCAGGACCC AGGTTTTGTA AGCGAAGACG
CCTCCTCCTT CACTAATAGC TCGTCCTGGG TCCAAAACAT TCGCTTCTGC

FIG.9A-28

23751 ACGAGGACCG CTCAGTACCA ACAGAGGATA AAAAGCAAGA CCAGGACAAC
TGCTCCTGGC GAGTCATGGT TGTCTCCTAT TTTTCGTTCT GGTCCTGTTG

23801 GCAGAGGCAA ACGAGGAACA AGTCGGGCGG GGGGACGAAA GGCATGGCGA
CGTCTCCGTT TGCTCCTTGT TCAGCCCGCC CCCCTGCTTT CCGTACCGCT

23851 CTACCTAGAT GTGGGAGACG ACGTGCTGTT GAAGCATCTG CAGCGCCAGT
GATGGATCTA CACCCTCTGC TGCACGACAA CTTCGTAGAC GTCGCGGTCA

23901 GCGCCATTAT CTGCGACGCG TTGCAAGAGC GCAGCGATGT GCCCCTCGCC
CGCGGTAATA GACGCTGCGC AACGTTCTCG CGTCGCTACA CGGGGAGCGG

23951 ATAGCGGATG TCAGCCTTGC CTACGAACGC CACCTATTCT CACCGCGCGT
TATCGCCTAC AGTCGGAACG GATGCTTGCG GTGGATAAGA GTGGCGCGCA

24001 ACCCCCCAAA CGCCAAGAAA ACGGCACATG CGAGCCCAAC CCGCGCCTCA
TGGGGGGTTT GCGGTTCTTT TGCCGTGTAC GCTCGGGTTG GGCGCGGAGT

24051 ACTTCTACCC CGTATTTGCC GTGCCAGAGG TGCTTGCCAC CTATCACATC
TGAAGATGGG GCATAAACGG CACGGTCTCC ACGAACGGTG GATAGTGTAG

24101 TTTTTCCAAA ACTGCAAGAT ACCCCTATCC TGCCGTGCCA ACCGCAGCCG
AAAAAGGTTT TGACGTTCTA TGGGGATAGG ACGGCACGGT TGGCGTCGGC

24151 AGCGGACAAG CAGCTGGCCT TGCGGCAGGG CGCTGTCATA CCTGATATCG
TCGCCTGTTC GTCGACCGGA ACGCCGTCCC GCGACAGTAT GGACTATAGC

24201 CCTCGCTCAA CGAAGTGCCA AAAATCTTTG AGGGTCTTGG ACGCGACGAG
GGAGCGAGTT GCTTCACGGT TTTTAGAAAC TCCCAGAACC TGCGCTGCTC

24251 AAGCGCGCGG CAAACGCTCT GCAACAGGAA AACAGCGAAA ATGAAAGTCA
TTCGCGCGCC GTTTGCGAGA CGTTGTCCTT TTGTCGCTTT TACTTTCAGT

24301 CTCTGGAGTG TTGGTGGAAC TCGAGGGTGA CAACGCGCGC CTAGCCGTAC
GAGACCTCAC AACCACCTTG AGCTCCCACT GTTGCGCGCG GATCGGCATG

24351 TAAAACGCAG CATCGAGGTC ACCCACTTTG CCTACCCGGC ACTTAACCTA
ATTTTGCGTC GTAGCTCCAG TGGGTGAAAC GGATGGGCCG TGAATTGGAT

24401 CCCCCCAAGG TCATGAGCAC AGTCATGAGT GAGCTGATCG TGCGCCGTGC
GGGGGGTTCC AGTACTCGTG TCAGTACTCA CTCGACTAGC ACGCGGCACG

24451 GCAGCCCCTG GAGAGGGATG CAAATTTGCA AGAACAAACA GAGGAGGGCC
CGTCGGGGAC CTCTCCCTAC GTTTAAACGT TCTTGTTTGT CTCCTCCCGG

24501 TACCCGCAGT TGGCGACGAG CAGCTAGCGC GCTGGCTTCA AACGCGCGAG
ATGGGCGTCA ACCGCTGCTC GTCGATCGCG CGACCGAAGT TTGCGCGCTC

24551 CCTGCCGACT TGGAGGAGCG ACGCAAACTA ATGATGGCCG CAGTGCTCGT
GGACGGCTGA ACCTCCTCGC TGCGTTTGAT TACTACCGGC GTCACGAGCA

FIG.9A-29

24601 TACCGTGGAG CTTGAGTGCA TGCAGCGGTT CTTTGCTGAC CCGGAGATGC
ATGGCACCTC GAACTCACGT ACGTCGCCAA GAAACGACTG GGCCTCTACG

24651 AGCGCAAGCT AGAGGAAACA TTGCACTACA CCTTTCGACA GGGCTACGTA
TCGCGTTCGA TCTCCTTTGT AACGTGATGT GGAAAGCTGT CCCGATGCAT

24701 CGCCAGGCCT GCAAGATCTC CAACGTGGAG CTCTGCAACC TGGTCTCCTA
GCGGTCCGGA CGTTCTAGAG GTTGCACCTC GAGACGTTGG ACCAGAGGAT

24751 CCTTGGAATT TTGCACGAAA ACCGCCTTGG GCAAAACGTG CTTCATTCCA
GGAACCTTAA AACGTGCTTT TGGCGGAACC CGTTTTGCAC GAAGTAAGGT

24801 CGCTCAAGGG CGAGGCGCGC CGCGACTACG TCCGCGACTG CGTTTACTTA
GCGAGTTCCC GCTCCGCGCG GCGCTGATGC AGGCGCTGAC GCAAATGAAT

24851 TTTCTATGCT ACACCTGGCA GACGGCCATG GGCGTTTGGC AGCAGTGCTT
AAAGATACGA TGTGGACCGT CTGCCGGTAC CCGCAAACCG TCGTCACGAA

24901 GGAGGAGTGC AACCTCAAGG AGCTGCAGAA ACTGCTAAAG CAAAACTTGA
CCTCCTCACG TTGGAGTTCC TCGACGTCTT TGACGATTTC GTTTTGAACT

24951 AGGACCTATG GACGGCCTTC AACGAGCGCT CCGTGGCCGC GCACCTGGCG
TCCTGGATAC CTGCCGGAAG TTGCTCGCGA GGCACCGGCG CGTGGACCGC

25001 GACATCATTT TCCCCGAACG CCTGCTTAAA ACCCTGCAAC AGGGTCTGCC
CTGTAGTAAA AGGGGCTTGC GGACGAATTT TGGGACGTTG TCCCAGACGG

25051 AGACTTCACC AGTCAAAGCA TGTTGCAGAA CTTTAGGAAC TTTATCCTAG
TCTGAAGTGG TCAGTTTCGT ACAACGTCTT GAAATCCTTG AAATAGGATC

25101 AGCGCTCAGG AATCTTGCCC GCCACCTGCT GTGCACTTCC TAGCGACTTT
TCGCGAGTCC TTAGAACGGG CGGTGGACGA CACGTGAAGG ATCGCTGAAA

25151 GTGCCCATTA AGTACCGCGA ATGCCCTCCG CCGCTTTGGG GCCACTGCTA
CACGGGTAAT TCATGGCGCT TACGGGAGGC GGCGAAACCC CGGTGACGAT

25201 CCTTCTGCAG CTAGCCAACT ACCTTGCCTA CCACTCTGAC ATAATGGAAG
GGAAGACGTC GATCGGTTGA TGGAACGGAT GGTGAGACTG TATTACCTTC

25251 ACGTGAGCGG TGACGGTCTA CTGGAGTGTC ACTGTCGCTG CAACCTATGC
TGCACTCGCC ACTGCCAGAT GACCTCACAG TGACAGCGAC GTTGGATACG

25301 ACCCCGCACC GCTCCCTGGT TTGCAATTCG CAGCTGCTTA ACGAAAGTCA
TGGGGCGTGG CGAGGGACCA AACGTTAAGC GTCGACGAAT TGCTTTCAGT

25351 AATTATCGGT ACCTTTGAGC TGCAGGGTCC CTCGCCTGAC GAAAAGTCCG
TTAATAGCCA TGGAAACTCG ACGTCCCAGG GAGCGGACTG CTTTTCAGGC

25401 CGGCTCCGGG GTTGAAACTC ACTCCGGGGC TGTGGACGTC GGCTTACCTT
GCCGAGGCCC CAACTTTGAG TGAGGCCCCG ACACCTGCAG CCGAATGGAA

FIG.9A-30

25451 CGCAAATTTG TACCTGAGGA CTACCACGCC CACGAGATTA GGTTCTACGA
GCGTTTAAAC ATGGACTCCT GATGGTGCGG GTGCTCTAAT CCAAGATGCT

25501 AGACCAATCC CGCCCGCCTA ATGCGGAGCT TACCGCCTGC GTCATTACCC
TCTGGTTAGG GCGGGCGGAT TACGCCTCGA ATGGCGGACG CAGTAATGGG

25551 AGGGCCACAT TCTTGGCCAA TTGCAAGCCA TCAACAAAGC CCGCCAAGAG
TCCCGGTGTA AGAACCGGTT AACGTTCGGT AGTTGTTTCG GGCGGTTCTC

25601 TTTCTGCTAC GAAAGGGACG GGGGGTTTAC TTGGACCCCC AGTCCGGCGA
AAAGACGATG CTTTCCCTGC CCCCCAAATG AACCTGGGGG TCAGGCCGCT

25651 GGAGCTCAAC CCAATCCCCC CGCCGCCGCA GCCCTATCAG CAGCAGCCGC
CCTCGAGTTG GGTTAGGGGG GCGGCGGCGT CGGGATAGTC GTCGTCGGCG

25701 GGGCCCTTGC TTCCCAGGAT GGCACCCAAA AAGAAGCTGC AGCTGCCGCC
CCCGGGAACG AAGGGTCCTA CCGTGGGTTT TTCTTCGACG TCGACGGCGG

25751 GCCACCCACG GACGAGGAGG AATACTGGGA CAGTCAGGCA GAGGAGGTTT
CGGTGGGTGC CTGCTCCTCC TTATGACCCT GTCAGTCCGT CTCCTCCAAA

25801 TGGACGAGGA GGAGGAGGAC ATGATGGAAG ACTGGGAGAG CCTAGACGAG
ACCTGCTCCT CCTCCTCCTG TACTACCTTC TGACCCTCTC GGATCTGCTC

25851 GAAGCTTCCG AGGTCGAAGA GGTGTCAGAC GAAACACCGT CACCCTCGGT
CTTCGAAGGC TCCAGCTTCT CCACAGTCTG CTTTGTGGCA GTGGGAGCCA

25901 CGCATTCCCC TCGCCGGCGC CCCAGAAATC GGCAACCGGT TCCAGCATGG
GCGTAAGGGG AGCGGCCGCG GGGTCTTTAG CCGTTGGCCA AGGTCGTACC

25951 CTACAACCTC CGCTCCTCAG GCGCCGCCGG CACTGCCCGT TCGCCGACCC
GATGTTGGAG GCGAGGAGTC CGCGGCGGCC GTGACGGGCA AGCGGCTGGG

26001 AACCGTAGAT GGGACACCAC TGGAACCAGG GCCGGTAAGT CCAAGCAGCC
TTGGCATCTA CCCTGTGGTG ACCTTGGTCC CGGCCATTCA GGTTCGTCGG

26051 GCCGCCGTTA GCCCAAGAGC AACAACAGCG CCAAGGCTAC CGCTCATGGC
CGGCGGCAAT CGGGTTCTCG TTGTTGTCGC GGTTCCGATG GCGAGTACCG

26101 GCGGGCACAA GAACGCCATA GTTGCTTGCT TGCAAGACTG TGGGGGCAAC
CGCCCGTGTT CTTGCGGTAT CAACGAACGA ACGTTCTGAC ACCCCCGTTG

26151 ATCTCCTTCG CCCGCCGCTT TCTTCTCTAC CATCACGGCG TGGCCTTCCC
TAGAGGAAGC GGGCGGCGAA AGAAGAGATG GTAGTGCCGC ACCGGAAGGG

26201 CCGTAACATC CTGCATTACT ACCGTCATCT CTACAGCCCA TACTGCACCG
GGCATTGTAG GACGTAATGA TGGCAGTAGA GATGTCGGGT ATGACGTGGC

26251 GCGGCAGCGG CAGCAACAGC AGCGGCCACA CAGAAGCAAA GGCGACCGGA
CGCCGTCGCC GTCGTTGTCG TCGCCGGTGT GTCTTCGTTT CCGCTGGCCT

FIG.9A-31

```
26301  TAGCAAGACT CTGACAAAGC CCAAGAAATC CACAGCGGCG GCAGCAGCAG
       ATCGTTCTGA GACTGTTTCG GGTTCTTTAG GTGTCGCCGC CGTCGTCGTC

26351  GAGGAGGAGC GCTGCGTCTG GCGCCCAACG AACCCGTATC GACCCGCGAG
       CTCCTCCTCG CGACGCAGAC CGCGGGTTGC TTGGGCATAG CTGGGCGCTC

26401  CTTAGAAACA GGATTTTTCC CACTCTGTAT GCTATATTTC AACAGAGCAG
       GAATCTTTGT CCTAAAAAGG GTGAGACATA CGATATAAAG TTGTCTCGTC

26451  GGGCCAAGAA CAAGAGCTGA AAATAAAAAA CAGGTCTCTG CGATCCCTCA
       CCCGGTTCTT GTTCTCGACT TTTATTTTTT GTCCAGAGAC GCTAGGGAGT

26501  CCCGCAGCTG CCTGTATCAC AAAAGCGAAG ATCAGCTTCG GCGCACGCTG
       GGGCGTCGAC GGACATAGTG TTTTCGCTTC TAGTCGAAGC CGCGTGCGAC

26551  GAAGACGCGG AGGCTCTCTT CAGTAAATAC TGCGCGCTGA CTCTTAAGGA
       CTTCTGCGCC TCCGAGAGAA GTCATTTATG ACGCGCGACT GAGAATTCCT

26601  CTAGTTTCGC GCCCTTTCTC AAATTTAAGC GCGAAAACTA CGTCATCTCC
       GATCAAAGCG CGGGAAAGAG TTTAAATTCG CGCTTTTGAT GCAGTAGAGG

26651  AGCGGCCACA CCCGGCGCCA GCACCTGTTG TCAGCGCCAT TATGAGCAAG
       TCGCCGGTGT GGGCCGCGGT CGTGGACAAC AGTCGCGGTA ATACTCGTTC

26701  GAAATTCCCA CGCCCTACAT GTGGAGTTAC CAGCCACAAA TGGGACTTGC
       CTTTAAGGGT GCGGGATGTA CACCTCAATG GTCGGTGTTT ACCCTGAACG

26751  GGCTGGAGCT GCCCAAGACT ACTCAACCCG AATAAACTAC ATGAGCGCGG
       CCGACCTCGA CGGGTTCTGA TGAGTTGGGC TTATTTGATG TACTCGCGCC

26801  GACCCCACAT GATATCCCGG GTCAACGGAA TACGCGCCCA CCGAAACCGA
       CTGGGGTGTA CTATAGGGCC CAGTTGCCTT ATGCGCGGGT GGCTTTGGCT

26851  ATTCTCCTGG AACAGGCGGC TATTACCACC ACACCTCGTA ATAACCTTAA
       TAAGAGGACC TTGTCCGCCG ATAATGGTGG TGTGGAGCAT TATTGGAATT

26901  TCCCCGTAGT TGGCCCGCTG CCCTGGTGTA CCAGGAAAGT CCCGCTCCCA
       AGGGGCATCA ACCGGGCGAC GGGACCACAT GGTCCTTTCA GGGCGAGGGT

26951  CCACTGTGGT ACTTCCCAGA GACGCCCAGG CCGAAGTTCA GATGACTAAC
       GGTGACACCA TGAAGGGTCT CTGCGGGTCC GGCTTCAAGT CTACTGATTG

27001  TCAGGGGCGC AGCTTGCGGG CGGCTTTCGT CACAGGGTGC GGTCGCCCGG
       AGTCCCCGCG TCGAACGCCC GCCGAAAGCA GTGTCCCACG CCAGCGGGCC

27051  GCAGGGTATA ACTCACCTGA CAATCAGAGG GCGAGGTATT CAGCTCAACG
       CGTCCCATAT TGAGTGGACT GTTAGTCTCC CGCTCCATAA GTCGAGTTGC

27101  ACGAGTCGGT GAGCTCCTCG CTTGGTCTCC GTCCGGACGG GACATTTCAG
       TGCTCAGCCA CTCGAGGAGC GAACCAGAGG CAGGCCTGCC CTGTAAAGTC
```

FIG.9A-32

27151 ATCGGCGGCG CCGGCCGCTC TTCATTCACG CCTCGTCAGG CAATCCTAAC
TAGCCGCCGC GGCCGGCGAG AAGTAAGTGC GGAGCAGTCC GTTAGGATTG

27201 TCTGCAGACC TCGTCCTCTG AGCCGCGCTC TGGAGGCATT GGAACTCTGC
AGACGTCTGG AGCAGGAGAC TCGGCGCGAG ACCTCCGTAA CCTTGAGACG

27251 AATTTATTGA GGAGTTTGTG CCATCGGTCT ACTTTAACCC CTTCTCGGGA
TTAAATAACT CCTCAAACAC GGTAGCCAGA TGAAATTGGG GAAGAGCCCT

27301 CCTCCCGGCC ACTATCCGGA TCAATTTATT CCTAACTTTG ACGCGGTAAA
GGAGGGCCGG TGATAGGCCT AGTTAAATAA GGATTGAAAC TGCGCCATTT

27351 GGACTCGGCG GACGGCTACG ACTGAATGTT AAGTGGAGAG GCAGAGCAAC
CCTGAGCCGC CTGCCGATGC TGACTTACAA TTCACCTCTC CGTCTCGTTG

27401 TGCGCCTGAA ACACCTGGTC CACTGTCGCC GCCACAAGTG CTTTGCCCGC
ACGCGGACTT TGTGGACCAG GTGACAGCGG CGGTGTTCAC GAAACGGGCG

27451 GACTCCGGTG AGTTTTGCTA CTTTGAATTG CCCGAGGATC ATATCGAGGG
CTGAGGCCAC TCAAAACGAT GAAACTTAAC GGGCTCCTAG TATAGCTCCC

27501 CCCGGCGCAC GGCGTCCGGC TTACCGCCCA GGGAGAGCTT GCCCGTAGCC
GGGCCGCGTG CCGCAGGCCG AATGGCGGGT CCCTCTCGAA CGGGCATCGG

27551 TGATTCGGGA GTTTACCCAG CGCCCCCTGC TAGTTGAGCG GGACAGGGGA
ACTAAGCCCT CAAATGGGTC GCGGGGGACG ATCAACTCGC CCTGTCCCCT

27601 CCCTGTGTTC TCACTGTGAT TTGCAACTGT CCTAACCCTG GATTACATCA
GGGACACAAG AGTGACACTA AACGTTGACA GGATTGGGAC CTAATGTAGT

27651 AGATCTTTGT TGCCATCTCT GTGCTGAGTA TAATAAATAC AGAAATTAAA
TCTAGAAACA ACGGTAGAGA CACGACTCAT ATTATTTATG TCTTTAATTT

27701 ATATACTGGG GCTCCTATCG CCATCCTGTA AACGCCACCG TCTTCACCCG
TATATGACCC CGAGGATAGC GGTAGGACAT TTGCGGTGGC AGAAGTGGGC

27751 CCCAAGCAAA CCAAGGCGAA CCTTACCTGG TACTTTTAAC ATCTCTCCCT
GGGTTCGTTT GGTTCCGCTT GGAATGGACC ATGAAAATTG TAGAGAGGGA

27801 CTGTGATTTA CAACAGTTTC AACCCAGACG GAGTGAGTCT ACGAGAGAAC
GACACTAAAT GTTGTCAAAG TTGGGTCTGC CTCACTCAGA TGCTCTCTTG

27851 CTCTCCGAGC TCAGCTACTC CATCAGAAAA AACACCACCC TCCTTACCTG
GAGAGGCTCG AGTCGATGAG GTAGTCTTTT TTGTGGTGGG AGGAATGGAC

27901 CCGGGAACGT ACGAGTGCGT CACCGGCCGC TGCACCACAC CTACCGCCTG
GGCCCTTGCA TGCTCACGCA GTGGCCGGCG ACGTGGTGTG GATGGCGGAC

27951 ACCGTAAACC AGACTTTTTC CGGACAGACC TCAATAACTC TGTTTACCAG
TGGCATTTGG TCTGAAAAAG GCCTGTCTGG AGTTATTGAG ACAAATGGTC

FIG.9A-33

28001 AACAGGAGGT GAGCTTAGAA AACCCTTAGG GTATTAGGCC AAAGGCGCAG
TTGTCCTCCA CTCGAATCTT TTGGGAATCC CATAATCCGG TTTCCGCGTC

28051 CTACTGTGGG GTTTATGAAC AATTCAAGCA ACTCTACGGG CTATTCTAAT
GATGACACCC CAAATACTTG TTAAGTTCGT TGAGATGCCC GATAAGATTA

28101 TCAGGTTTCT CTAGAATCGG GGTTGGGGTT ATTCTCTGTC TTGTGATTCT
AGTCCAAAGA GATCTTAGCC CCAACCCCAA TAAGAGACAG AACACTAAGA

28151 CTTTATTCTT ATACTAACGC TTCTCTGCCT AAGGCTCGCC GCCTGCTGTG
GAAATAAGAA TATGATTGCG AAGAGACGGA TTCCGAGCGG CGGACGACAC

28201 TGCACATTTG CATTTATTGT CAGCTTTTTA AACGCTGGGG TCGCCACCCA
ACGTGTAAAC GTAAATAACA GTCGAAAAAT TTGCGACCCC AGCGGTGGGT

28251 AGATGATTAG GTACATAATC CTAGGTTTAC TCACCCTTGC GTCAGCCCAC
TCTACTAATC CATGTATTAG GATCCAAATG AGTGGGAACG CAGTCGGGTG

28301 GGTACCACCC AAAAGGTGGA TTTTAAGGAG CCAGCCTGTA ATGTTACATT
CCATGGTGGG TTTTCCACCT AAAATTCCTC GGTCGGACAT TACAATGTAA

28351 CGCAGCTGAA GCTAATGAGT GCACCACTCT TATAAAATGC ACCACAGAAC
GCGTCGACTT CGATTACTCA CGTGGTGAGA ATATTTTACG TGGTGTCTTG

28401 ATGAAAAGCT GCTTATTCGC CACAAAAACA AAATTGGCAA GTATGCTGTT
TACTTTTCGA CGAATAAGCG GTGTTTTTGT TTTAACCGTT CATACGACAA

28451 TATGCTATTT GGCAGCCAGG TGACACTACA GAGTATAATG TTACAGTTTT
ATACGATAAA CCGTCGGTCC ACTGTGATGT CTCATATTAC AATGTCAAAA

28501 CCAGGGTAAA AGTCATAAAA CTTTTATGTA TACTTTTCCA TTTTATGAAA
GGTCCCATTT TCAGTATTTT GAAAATACAT ATGAAAAGGT AAAATACTTT

28551 TGTGCGACAT TACCATGTAC ATGAGCAAAC AGTATAAGTT GTGGCCCCCA
ACACGCTGTA ATGGTACATG TACTCGTTTG TCATATTCAA CACCGGGGGT

28601 CAAAATTGTG TGGAAAACAC TGGCACTTTC TGCTGCACTG CTATGCTAAT
GTTTTAACAC ACCTTTTGTG ACCGTGAAAG ACGACGTGAC GATACGATTA

28651 TACAGTGCTC GCTTTGGTCT GTACCCTACT CTATATTAAA TACAAAAGCA
ATGTCACGAG CGAAACCAGA CATGGGATGA GATATAATTT ATGTTTTCGT

28701 GACGCAGCTT TATTGAGGAA AAGAAAATGC CTTAATTTAC TAAGTTACAA
CTGCGTCGAA ATAACTCCTT TTCTTTTACG GAATTAAATG ATTCAATGTT

28751 AGCTAATGTC ACCACTAACT GCTTTACTCG CTGCTTGCAA AACAAATTCA
TCGATTACAG TGGTGATTGA CGAAATGAGC GACGAACGTT TTGTTTAAGT

28801 AAAAGTTAGC ATTATAATTA GAATAGGATT TAAACCCCCC GGTCATTTCC
TTTTCAATCG TAATATTAAT CTTATCCTAA ATTTGGGGGG CCAGTAAAGG

FIG.9A-34

28851 TGCTCAATAC CATTCCCCTG AACAATTGAC TCTATGTGGG ATATGCTCCA
ACGAGTTATG GTAAGGGGAC TTGTTAACTG AGATACACCC TATACGAGGT

28901 GCGCTACAAC CTTGAAGTCA GGCTTCCTGG ATGTCAGCAT CTGACTTTGG
CGCGATGTTG GAACTTCAGT CCGAAGGACC TACAGTCGTA GACTGAAACC

28951 CCAGCACCTG TCCCGCGGAT TTGTTCCAGT CCAACTACAG CGACCCACCC
GGTCGTGGAC AGGGCGCCTA AACAAGGTCA GGTTGATGTC GCTGGGTGGG

29001 TAACAGAGAT GACCAACACA ACCAACGCGG CCGCCGCTAC CGGACTTACA
ATTGTCTCTA CTGGTTGTGT TGGTTGCGCC GGCGGCGATG GCCTGAATGT

29051 TCTACCACAA ATACACCCCA AGTTTCTGCC TTTGTCAATA ACTGGGATAA
AGATGGTGTT TATGTGGGGT TCAAAGACGG AAACAGTTAT TGACCCTATT

29101 CTTGGGCATG TGGTGGTTCT CCATAGCGCT TATGTTTGTA TGCCTTATTA
GAACCCGTAC ACCACCAAGA GGTATCGCGA ATACAAACAT ACGGAATAAT

29151 TTATGTGGCT CATCTGCTGC CTAAAGCGCA AACGCGCCCG ACCACCCATC
AATACACCGA GTAGACGACG GATTTCGCGT TTGCGCGGGC TGGTGGGTAG

29201 TATAGTCCCA TCATTGTGCT ACACCCAAAC AATGATGGAA TCCATAGATT
ATATCAGGGT AGTAACACGA TGTGGGTTTG TTACTACCTT AGGTATCTAA

29251 GGACGGACTG AAACACATGT TCTTTTCTCT TACAGTATGA TTAAATGAGA
CCTGCCTGAC TTTGTGTACA AGAAAAGAGA ATGTCATACT AATTTACTCT

29301 CATGATTCCT CGAGTTTTTA TATTACTGAC CCTTGTTGCG CTTTTTTGTG
GTACTAAGGA GCTCAAAAAT ATAATGACTG GGAACAACGC GAAAAAACAC

29351 CGTGCTCCAC ATTGGCTGCG GTTTCTCACA TCGAAGTAGA CTGCATTCCA
GCACGAGGTG TAACCGACGC CAAAGAGTGT AGCTTCATCT GACGTAAGGT

29401 GCCTTCACAG TCTATTTGCT TTACGGATTT GTCACCCTCA CGCTCATCTG
CGGAAGTGTC AGATAAACGA AATGCCTAAA CAGTGGGAGT GCGAGTAGAC

29451 CAGCCTCATC ACTGTGGTCA TCGCCTTTAT CCAGTGCATT GACTGGGTCT
GTCGGAGTAG TGACACCAGT AGCGGAAATA GGTCACGTAA CTGACCCAGA

29501 GTGTGCGCTT TGCATATCTC AGACACCATC CCCAGTACAG GGACAGGACT
CACACGCGAA ACGTATAGAG TCTGTGGTAG GGGTCATGTC CCTGTCCTGA

29551 ATAGCTGAGC TTCTTAGAAT TCTTTAATTA TGAAATTTAC TGTGACTTTT
TATCGACTCG AAGAATCTTA AGAAATTAAT ACTTTAAATG ACACTGAAAA

29601 CTGCTGATTA TTTGCACCCT ATCTGCGTTT TGTTCCCCGA CCTCCAAGCC
GACGACTAAT AAACGTGGGA TAGACGCAAA ACAAGGGGCT GGAGGTTCGG

29651 TCAAAGACAT ATATCATGCA GATTCACTCG TATATGGAAT ATTCCAAGTT
AGTTTCTGTA TATAGTACGT CTAAGTGAGC ATATACCTTA TAAGGTTCAA

FIG.9A-35

```
29701  GCTACAATGA AAAAAGCGAT CTTTCCGAAG CCTGGTTATA TGCAATCATC
       CGATGTTACT TTTTTCGCTA GAAAGGCTTC GGACCAATAT ACGTTAGTAG

29751  TCTGTTATGG TGTTCTGCAG TACCATCTTA GCCCTAGCTA TATATCCCTA
       AGACAATACC ACAAGACGTC ATGGTAGAAT CGGGATCGAT ATATAGGGAT

29801  CCTTGACATT GGCTGGAACG CAATAGATGC CATGAACCAC CCAACTTTCC
       GGAACTGTAA CCGACCTTGC GTTATCTACG GTACTTGGTG GGTTGAAAGG

29851  CCGCGCCCGC TATGCTTCCA CTGCAACAAG TTGTTGCCGG CGGCTTTGTC
       GGCGCGGGCG ATACGAAGGT GACGTTGTTC AACAACGGCC GCCGAAACAG

29901  CCAGCCAATC AGCCTCGCCC ACCTTCTCCC ACCCCCACTG AAATCAGCTA
       GGTCGGTTAG TCGGAGCGGG TGGAAGAGGG TGGGGGTGAC TTTAGTCGAT

29951  CTTTAATCTA ACAGGAGGAG ATGACTGACA CCCTAGATCT AGAAATGGAC
       GAAATTAGAT TGTCCTCCTC TACTGACTGT GGGATCTAGA TCTTTACCTG

30001  GGAATTATTA CAGAGCAGCG CCTGCTAGAA AGACGCAGGG CAGCGGCCGA
       CCTTAATAAT GTCTCGTCGC GGACGATCTT TCTGCGTCCC GTCGCCGGCT

30051  GCAACAGCGC ATGAATCAAG AGCTCCAAGA CATGGTTAAC TTGCACCAGT
       CGTTGTCGCG TACTTAGTTC TCGAGGTTCT GTACCAATTG AACGTGGTCA

30101  GCAAAAGGGG TATCTTTTGT CTCGTAAAGC AGGCCAAAGT CACCTACGAC
       CGTTTTCCCC ATAGAAAACA GAGCATTTCG TCCGGTTTCA GTGGATGCTG

30151  AGTAATACCA CCGGACACCG CCTTAGCTAC AAGTTGCCAA CCAAGCGTCA
       TCATTATGGT GGCCTGTGGC GGAATCGATG TTCAACGGTT GGTTCGCAGT

30201  GAAATTGGTG GTCATGGTGG GAGAAAAGCC CATTACCATA ACTCAGCACT
       CTTTAACCAC CAGTACCACC CTCTTTTCGG GTAATGGTAT TGAGTCGTGA

30251  CGGTAGAAAC CGAAGGCTGC ATTCACTCAC CTTGTCAAGG ACCTGAGGAT
       GCCATCTTTG GCTTCCGACG TAAGTGAGTG GAACAGTTCC TGGACTCCTA

30301  CTCTGCACCC TTATTAAGAC CCTGTGCGGT CTCAAAGATC TTATTCCCTT
       GAGACGTGGG AATAATTCTG GGACACGCCA GAGTTTCTAG AATAAGGGAA

30351  TAACTAATAA AAAAAAATAA TAAAGCATCA CTTACTTAAA ATCAGTTAGC
       ATTGATTATT TTTTTTTATT ATTTCGTAGT GAATGAATTT TAGTCAATCG

30401  AAATTTCTGT CCAGTTTATT CAGCAGCACC TCCTTGCCCT CCTCCCAGCT
       TTTAAAGACA GGTCAAATAA GTCGTCGTGG AGGAACGGGA GGAGGGTCGA

30451  CTGGTATTGC AGCTTCCTCC TGGCTGCAAA CTTTCTCCAC AATCTAAATG
       GACCATAACG TCGAAGGAGG ACCGACGTTT GAAAGAGGTG TTAGATTTAC

30501  GAATGTCAGT TTCCTCCTGT TCCTGTCCAT CCGCACCCAC TATCTTCATG
       CTTACAGTCA AAGGAGGACA AGGACAGGTA GGCGTGGGTG ATAGAAGTAC
```

FIG.9A-36

30551 TTGTTGCAGA TGAAGCGCGC AAGACCGTCT GAAGATACCT TCAACCCCGT
AACAACGTCT ACTTCGCGCG TTCTGGCAGA CTTCTATGGA AGTTGGGGCA

30601 GTATCCATAT GACACGGAAA CCGGTCCTCC AACTGTGCCT TTTCTTACTC
CATAGGTATA CTGTGCCTTT GGCCAGGAGG TTGACACGGA AAAGAATGAG

30651 CTCCCTTTGT ATCCCCCAAT GGGTTTCAAG AGAGTCCCCC TGGGGTACTC
GAGGGAAACA TAGGGGGTTA CCCAAAGTTC TCTCAGGGGG ACCCCATGAG

30701 TCTTTGCGCC TATCCGAACC TCTAGTTACC TCCAATGGCA TGCTTGCGCT
AGAAACGCGG ATAGGCTTGG AGATCAATGG AGGTTACCGT ACGAACGCGA

30751 CAAAATGGGC AACGGCCTCT CTCTGGACGA GGCCGGCAAC CTTACCTCCC
GTTTTACCCG TTGCCGGAGA GAGACCTGCT CCGGCCGTTG GAATGGAGGG

30801 AAAATGTAAC CACTGTGAGC CCACCTCTCA AAAAAACCAA GTCAAACATA
TTTTACATTG GTGACACTCG GGTGGAGAGT TTTTTTGGTT CAGTTTGTAT

30851 AACCTGGAAA TATCTGCACC CCTCACAGTT ACCTCAGAAG CCCTAACTGT
TTGGACCTTT ATAGACGTGG GGAGTGTCAA TGGAGTCTTC GGGATTGACA

30901 GGCTGCCGCC GCACCTCTAA TGGTCGCGGG CAACACACTC ACCATGCAAT
CCGACGGCGG CGTGGAGATT ACCAGCGCCC GTTGTGTGAG TGGTACGTTA

30951 CACAGGCCCC GCTAACCGTG CACGACTCCA AACTTAGCAT TGCCACCCAA
GTGTCCGGGG CGATTGGCAC GTGCTGAGGT TTGAATCGTA ACGGTGGGTT

31001 GGACCCCTCA CAGTGTCAGA AGGAAAGCTA GCCCTGCAAA CATCAGGCCC
CCTGGGGAGT GTCACAGTCT TCCTTTCGAT CGGGACGTTT GTAGTCCGGG

31051 CCTCACCACC ACCGATAGCA GTACCCTTAC TATCACTGCC TCACCCCCTT
GGAGTGGTGG TGGCTATCGT CATGGGAATG ATAGTGACGG AGTGGGGGAA

31101 TAACTACTGC CACTGGTAGC TTGGGCATTG ACTTGAAAGA GCCCATTTAT
ATTGATGACG GTGACCATCG AACCCGTAAC TGAACTTTCT CGGGTAAATA

31151 ACACAAAATG GAAAACTAGG ACTAAAGTAC GGGGCTCCTT TGCATGTAAC
TGTGTTTTAC CTTTTGATCC TGATTTCATG CCCCGAGGAA ACGTACATTG

31201 AGACGACCTA AACACTTTGA CCGTAGCAAC TGGTCCAGGT GTGACTATTA
TCTGCTGGAT TTGTGAAACT GGCATCGTTG ACCAGGTCCA CACTGATAAT

31251 ATAATACTTC CTTGCAAACT AAAGTTACTG GAGCCTTGGG TTTTGATTCA
TATTATGAAG GAACGTTTGA TTTCAATGAC CTCGGAACCC AAAACTAAGT

31301 CAAGGCAATA TGCAACTTAA TGTAGCAGGA GGACTAAGGA TTGATTCTCA
GTTCCGTTAT ACGTTGAATT ACATCGTCCT CCTGATTCCT AACTAAGAGT

31351 AAACAGACGC CTTATACTTG ATGTTAGTTA TCCGTTTGAT GCTCAAAACC
TTTGTCTGCG GAATATGAAC TACAATCAAT AGGCAAACTA CGAGTTTTGG

FIG.9A-37

31401 AACTAAATCT AAGACTAGGA CAGGGCCCTC TTTTTATAAA CTCAGCCCAC
TTGATTTAGA TTCTGATCCT GTCCCGGGAG AAAAATATTT GAGTCGGGTG

31451 AACTTGGATA TTAACTACAA CAAAGGCCTT TACTTGTTTA CAGCTTCAAA
TTGAACCTAT AATTGATGTT GTTTCCGGAA ATGAACAAAT GTCGAAGTTT

31501 CAATTCCAAA AAGCTTGAGG TTAACCTAAG CACTGCCAAG GGGTTGATGT
GTTAAGGTTT TTCGAACTCC AATTGGATTC GTGACGGTTC CCCAACTACA

31551 TTGACGCTAC AGCCATAGCC ATTAATGCAG GAGATGGGCT TGAATTTGGT
AACTGCGATG TCGGTATCGG TAATTACGTC CTCTACCCGA ACTTAAACCA

31601 TCACCTAATG CACCAAACAC AAATCCCCTC AAAACAAAAA TTGGCCATGG
AGTGGATTAC GTGGTTTGTG TTTAGGGGAG TTTTGTTTTT AACCGGTACC

31651 CCTAGAATTT GATTCAAACA AGGCTATGGT TCCTAAACTA GGAACTGGCC
GGATCTTAAA CTAAGTTTGT TCCGATACCA AGGATTTGAT CCTTGACCGG

31701 TTAGTTTTGA CAGCACAGGT GCCATTACAG TAGGAAACAA AAATAATGAT
AATCAAAACT GTCGTGTCCA CGGTAATGTC ATCCTTTGTT TTTATTACTA

31751 AAGCTAACTT TGTGGACCAC ACCAGCTCCA TCTCCTAACT GTAGACTAAA
TTCGATTGAA ACACCTGGTG TGGTCGAGGT AGAGGATTGA CATCTGATTT

31801 TGCAGAGAAA GATGCTAAAC TCACTTTGGT CTTAACAAAA TGTGGCAGTC
ACGTCTCTTT CTACGATTTG AGTGAAACCA GAATTGTTTT ACACCGTCAG

31851 AAATACTTGC TACAGTTTCA GTTTTGGCTG TTAAAGGCAG TTTGGCTCCA
TTTATGAACG ATGTCAAAGT CAAAACCGAC AATTTCCGTC AAACCGAGGT

31901 ATATCTGGAA CAGTTCAAAG TGCTCATCTT ATTATAAGAT TTGACGAAAA
TATAGACCTT GTCAAGTTTC ACGAGTAGAA TAATATTCTA AACTGCTTTT

31951 TGGAGTGCTA CTAAACAATT CCTTCCTGGA CCCAGAATAT TGGAACTTTA
ACCTCACGAT GATTTGTTAA GGAAGGACCT GGGTCTTATA ACCTTGAAAT

32001 GAAATGGAGA TCTTACTGAA GGCACAGCCT ATACAAACGC TGTTGGATTT
CTTTACCTCT AGAATGACTT CCGTGTCGGA TATGTTTGCG ACAACCTAAA

32051 ATGCCTAACC TATCAGCTTA TCCAAAATCT CACGGTAAAA CTGCCAAAAG
TACGGATTGG ATAGTCGAAT AGGTTTTAGA GTGCCATTTT GACGGTTTTC

32101 TAACATTGTC AGTCAAGTTT ACTTAAACGG AGACAAAACT AAACCTGTAA
ATTGTAACAG TCAGTTCAAA TGAATTTGCC TCTGTTTTGA TTTGGACATT

32151 CACTAACCAT TACACTAAAC GGTACACAGG AAACAGGAGA CACAACTCCA
GTGATTGGTA ATGTGATTTG CCATGTGTCC TTTGTCCTCT GTGTTGAGGT

32201 AGTGCATACT CTATGTCATT TTCATGGGAC TGGTCTGGCC ACAACTACAT
TCACGTATGA GATACAGTAA AAGTACCCTG ACCAGACCGG TGTTGATGTA

FIG.9A-38

32251 TAATGAAATA TTTGCCACAT CCTCTTACAC TTTTTCATAC ATTGCCCAAG
ATTACTTTAT AAACGGTGTA GGAGAATGTG AAAAAGTATG TAACGGGTTC

32301 AATAAAGAAT CGTTTGTGTT ATGTTTCAAC GTGTTTATTT TTCAATTGCA
TTATTTCTTA GCAAACACAA TACAAAGTTG CACAAATAAA AAGTTAACGT

32351 GAAAATTTCA AGTCATTTTT CATTCAGTAG TATAGCCCCA CCACCACATA
CTTTTAAAGT TCAGTAAAAA GTAAGTCATC ATATCGGGGT GGTGGTGTAT

32401 GCTTATACAG ATCACCGTAC CTTAATCAAA CTCACAGAAC CCTAGTATTC
CGAATATGTC TAGTGGCATG GAATTAGTTT GAGTGTCTTG GGATCATAAG

32451 AACCTGCCAC CTCCCTCCCA ACACACAGAG TACACAGTCC TTTCTCCCCG
TTGGACGGTG GAGGGAGGGT TGTGTGTCTC ATGTGTCAGG AAAGAGGGGC

32501 GCTGGCCTTA AAAAGCATCA TATCATGGGT AACAGACATA TTCTTAGGTG
CGACCGGAAT TTTTCGTAGT ATAGTACCCA TTGTCTGTAT AAGAATCCAC

32551 TTATATTCCA CACGGTTTCC TGTCGAGCCA AACGCTCATC AGTGATATTA
AATATAAGGT GTGCCAAAGG ACAGCTCGGT TTGCGAGTAG TCACTATAAT

32601 ATAAACTCCC CGGGCAGCTC ACTTAAGTTC ATGTCGCTGT CCAGCTGCTG
TATTTGAGGG GCCCGTCGAG TGAATTCAAG TACAGCGACA GGTCGACGAC

32651 AGCCACAGGC TGCTGTCCAA CTTGCGGTTG CTTAACGGGC GGCGAAGGAG
TCGGTGTCCG ACGACAGGTT GAACGCCAAC GAATTGCCCG CCGCTTCCTC

32701 AAGTCCACGC CTACATGGGG GTAGAGTCAT AATCGTGCAT CAGGATAGGG
TTCAGGTGCG GATGTACCCC CATCTCAGTA TTAGCACGTA GTCCTATCCC

32751 CGGTGGTGCT GCAGCAGCGC GCGAATAAAC TGCTGCCGCC GCCGCTCCGT
GCCACCACGA CGTCGTCGCG CGCTTATTTG ACGACGGCGG CGGCGAGGCA

32801 CCTGCAGGAA TACAACATGG CAGTGGTCTC CTCAGCGATG ATTCGCACCG
GGACGTCCTT ATGTTGTACC GTCACCAGAG GAGTCGCTAC TAAGCGTGGC

32851 CCCGCAGCAT AAGGCGCCTT GTCCTCCGGG CACAGCAGCG CACCCTGATC
GGGCGTCGTA TTCCGCGGAA CAGGAGGCCC GTGTCGTCGC GTGGGACTAG

32901 TCACTTAAAT CAGCACAGTA ACTGCAGCAC AGCACCACAA TATTGTTCAA
AGTGAATTTA GTCGTGTCAT TGACGTCGTG TCGTGGTGTT ATAACAAGTT

32951 AATCCCACAG TGCAAGGCGC TGTATCCAAA GCTCATGGCG GGGACCACAG
TTAGGGTGTC ACGTTCCGCG ACATAGGTTT CGAGTACCGC CCCTGGTGTC

33001 AACCCACGTG GCCATCATAC CACAAGCGCA GGTAGATTAA GTGGCGACCC
TTGGGTGCAC CGGTAGTATG GTGTTCGCGT CCATCTAATT CACCGCTGGG

33051 CTCATAAACA CGCTGGACAT AAACATTACC TCTTTTGGCA TGTTGTAATT
GAGTATTTGT GCGACCTGTA TTTGTAATGG AGAAAACCGT ACAACATTAA

FIG.9A-39

33101 CACCACCTCC CGGTACCATA TAAACCTCTG ATTAAACATG GCGCCATCCA
GTGGTGGAGG GCCATGGTAT ATTTGGAGAC TAATTTGTAC CGCGGTAGGT

33151 CCACCATCCT AAACCAGCTG GCCAAAACCT GCCCGCCGGC TATACACTGC
GGTGGTAGGA TTTGGTCGAC CGGTTTTGGA CGGGCGGCCG ATATGTGACG

33201 AGGGAACCGG GACTGGAACA ATGACAGTGG AGAGCCCAGG ACTCGTAACC
TCCCTTGGCC CTGACCTTGT TACTGTCACC TCTCGGGTCC TGAGCATTGG

33251 ATGGATCATC ATGCTCGTCA TGATATCAAT GTTGGCACAA CACAGGCACA
TACCTAGTAG TACGAGCAGT ACTATAGTTA CAACCGTGTT GTGTCCGTGT

33301 CGTGCATACA CTTCCTCAGG ATTACAAGCT CCTCCCGCGT TAGAACCATA
GCACGTATGT GAAGGAGTCC TAATGTTCGA GGAGGGCGCA ATCTTGGTAT

33351 TCCCAGGGAA CAACCCATTC CTGAATCAGC GTAAATCCCA CACTGCAGGG
AGGGTCCCTT GTTGGGTAAG GACTTAGTCG CATTTAGGGT GTGACGTCCC

33401 AAGACCTCGC ACGTAACTCA CGTTGTGCAT TGTCAAAGTG TTACATTCGG
TTCTGGAGCG TGCATTGAGT GCAACACGTA ACAGTTTCAC AATGTAAGCC

33451 GCAGCAGCGG ATGATCCTCC AGTATGGTAG CGCGGGTTTC TGTCTCAAAA
CGTCGTCGCC TACTAGGAGG TCATACCATC GCGCCCAAAG ACAGAGTTTT

33501 GGAGGTAGAC GATCCCTACT GTACGGAGTG CGCCGAGACA ACCGAGATCG
CCTCCATCTG CTAGGGATGA CATGCCTCAC GCGGCTCTGT TGGCTCTAGC

33551 TGTTGGTCGT AGTGTCATGC CAAATGGAAC GCCGGACGTA GTCATATTTC
ACAACCAGCA TCACAGTACG GTTTACCTTG CGGCCTGCAT CAGTATAAAG

33601 CTGAAGCAAA ACCAGGTGCG GGCGTGACAA ACAGATCTGC GTCTCCGGTC
GACTTCGTTT TGGTCCACGC CCGCACTGTT TGTCTAGACG CAGAGGCCAG

33651 TCGCCGCTTA GATCGCTCTG TGTAGTAGTT GTAGTATATC CACTCTCTCA
AGCGGCGAAT CTAGCGAGAC ACATCATCAA CATCATATAG GTGAGAGAGT

33701 AAGCATCCAG GCGCCCCCTG GCTTCGGGTT CTATGTAAAC TCCTTCATGC
TTCGTAGGTC CGCGGGGGAC CGAAGCCCAA GATACATTTG AGGAAGTACG

33751 GCCGCTGCCC TGATAACATC CACCACCGCA GAATAAGCCA CACCCAGCCA
CGGCGACGGG ACTATTGTAG GTGGTGGCGT CTTATTCGGT GTGGGTCGGT

33801 ACCTACACAT TCGTTCTGCG AGTCACACAC GGGAGGAGCG GGAAGAGCTG
TGGATGTGTA AGCAAGACGC TCAGTGTGTG CCCTCCTCGC CCTTCTCGAC

33851 GAAGAACCAT GTTTTTTTTT TTATTCCAAA AGATTATCCA AAACCTCAAA
CTTCTTGGTA CAAAAAAAAA AATAAGGTTT TCTAATAGGT TTTGGAGTTT

33901 ATGAAGATCT ATTAAGTGAA CGCGCTCCCC TCCGGTGGCG TGGTCAAACT
TACTTCTAGA TAATTCACTT GCGCGAGGGG AGGCCACCGC ACCAGTTTGA

FIG.9A-40

33951 CTACAGCCAA AGAACAGATA ATGGCATTTG TAAGATGTTG CACAATGGCT
GATGTCGGTT TCTTGTCTAT TACCGTAAAC ATTCTACAAC GTGTTACCGA

34001 TCCAAAAGGC AAACGGCCCT CACGTCCAAG TGGACGTAAA GGCTAAACCC
AGGTTTTCCG TTTGCCGGGA GTGCAGGTTC ACCTGCATTT CCGATTTGGG

34051 TTCAGGGTGA ATCTCCTCTA TAAACATTCC AGCACCTTCA ACCATGCCCA
AAGTCCCACT TAGAGGAGAT ATTTGTAAGG TCGTGGAAGT TGGTACGGGT

34101 AATAATTCTC ATCTCGCCAC CTTCTCAATA TATCTCTAAG CAAATCCCGA
TTATTAAGAG TAGAGCGGTG GAAGAGTTAT ATAGAGATTC GTTTAGGGCT

34151 ATATTAAGTC CGGCCATTGT AAAAATCTGC TCCAGAGCGC CCTCCACCTT
TATAATTCAG GCCGGTAACA TTTTTAGACG AGGTCTCGCG GGAGGTGGAA

34201 CAGCCTCAAG CAGCGAATCA TGATTGCAAA AATTCAGGTT CCTCACAGAC
GTCGGAGTTC GTCGCTTAGT ACTAACGTTT TTAAGTCCAA GGAGTGTCTG

34251 CTGTATAAGA TTCAAAAGCG GAACATTAAC AAAAATACCG CGATCCCGTA
GACATATTCT AAGTTTTCGC CTTGTAATTG TTTTTATGGC GCTAGGGCAT

34301 GGTCCCTTCG CAGGGCCAGC TGAACATAAT CGTGCAGGTC TGCACGGACC
CCAGGGAAGC GTCCCGGTCG ACTTGTATTA GCACGTCCAG ACGTGCCTGG

34351 AGCGCGGCCA CTTCCCCGCC AGGAACCATG ACAAAAGAAC CCACACTGAT
TCGCGCCGGT GAAGGGGCGG TCCTTGGTAC TGTTTTCTTG GGTGTGACTA

34401 TATGACACGC ATACTCGGAG CTATGCTAAC CAGCGTAGCC CCGATGTAAG
ATACTGTGCG TATGAGCCTC GATACGATTG GTCGCATCGG GGCTACATTC

34451 CTTGTTGCAT GGGCGGCGAT ATAAAATGCA AGGTGCTGCT CAAAAAATCA
GAACAACGTA CCCGCCGCTA TATTTTACGT TCCACGACGA GTTTTTTAGT

34501 GGCAAAGCCT CGCGCAAAAA AGAAAGCACA TCGTAGTCAT GCTCATGCAG
CCGTTTCGGA GCGCGTTTTT TCTTTCGTGT AGCATCAGTA CGAGTACGTC

34551 ATAAAGGCAG GTAAGCTCCG GAACCACCAC AGAAAAAGAC ACCATTTTTC
TATTTCCGTC CATTCGAGGC CTTGGTGGTG TCTTTTTCTG TGGTAAAAAG

34601 TCTCAAACAT GTCTGCGGGT TTCTGCATAA ACACAAAATA AAATAACAAA
AGAGTTTGTA CAGACGCCCA AAGACGTATT TGTGTTTTAT TTTATTGTTT

34651 AAAACATTTA AACATTAGAA GCCTGTCTTA CAACAGGAAA AACAACCCTT
TTTTGTAAAT TTGTAATCTT CGGACAGAAT GTTGTCCTTT TTGTTGGGAA

34701 ATAAGCATAA GACGGACTAC GGCCATGCCG GCGTGACCGT AAAAAAACTG
TATTCGTATT CTGCCTGATG CCGGTACGGC CGCACTGGCA TTTTTTTGAC

34751 GTCACCGTGA TTAAAAAGCA CCACCGACAG CTCCTCGGTC ATGTCCGGAG
CAGTGGCACT AATTTTTCGT GGTGGCTGTC GAGGAGCCAG TACAGGCCTC

FIG.9A-41

```
34801  TCATAATGTA AGACTCGGTA AACACATCAG GTTGATTCAC ATCGGTCAGT
       AGTATTACAT TCTGAGCCAT TTGTGTAGTC CAACTAAGTG TAGCCAGTCA

34851  GCTAAAAAGC GACCGAAATA GCCCGGGGGA ATACATACCC GCAGGCGTAG
       CGATTTTTCG CTGGCTTTAT CGGGCCCCCT TATGTATGGG CGTCCGCATC

34901  AGACAACATT ACAGCCCCCA TAGGAGGTAT AACAAAATTA ATAGGAGAGA
       TCTGTTGTAA TGTCGGGGGT ATCCTCCATA TTGTTTTAAT TATCCTCTCT

34951  AAAACACATA AACACCTGAA AAACCCTCCT GCCTAGGCAA AATAGCACCC
       TTTTGTGTAT TTGTGGACTT TTTGGGAGGA CGGATCCGTT TTATCGTGGG

35001  TCCCGCTCCA GAACAACATA CAGCGCTTCC ACAGCGGCAG CCATAACAGT
       AGGGCGAGGT CTTGTTGTAT GTCGCGAAGG TGTCGCCGTC GGTATTGTCA

35051  CAGCCTTACC AGTAAAAAAG AAAACCTATT AAAAAAACAC CACTCGACAC
       GTCGGAATGG TCATTTTTTC TTTTGGATAA TTTTTTTGTG GTGAGCTGTG

35101  GGCACCAGCT CAATCAGTCA CAGTGTAAAA AAGGGCCAAG TGCAGAGCGA
       CCGTGGTCGA GTTAGTCAGT GTCACATTTT TTCCCGGTTC ACGTCTCGCT

35151  GTATATATAG GACTAAAAAA TGACGTAACG GTTAAAGTCC ACAAAAAACA
       CATATATATC CTGATTTTTT ACTGCATTGC CAATTTCAGG TGTTTTTTGT

35201  CCCAGAAAAC CGCACGCGAA CCTACGCCCA GAAACGAAAG CCAAAAAACC
       GGGTCTTTTG GCGTGCGCTT GGATGCGGGT CTTTGCTTTC GGTTTTTTGG

35251  CACAACTTCC TCAAATCGTC ACTTCCGTTT TCCCACGTTA CGTCACTTCC
       GTGTTGAAGG AGTTTAGCAG TGAAGGCAAA AGGGTGCAAT GCAGTGAAGG

35301  CATTTTAAGA AAACTACAAT TCCCAACACA TACAAGTTAC TCCGCCCTAA
       GTAAAATTCT TTTGATGTTA AGGGTTGTGT ATGTTCAATG AGGCGGGATT

35351  AACCTACGTC ACCCGCCCCG TTCCCACGCC CCGCGCCACG TCACAAACTC
       TTGGATGCAG TGGGCGGGGC AAGGGTGCGG GGCGCGGTGC AGTGTTTGAG

35401  CACCCCCTCA TTATCATATT GGCTTCAATC CAAAATAAGG TATATTATTG
       GTGGGGGAGT AATAGTATAA CCGAAGTTAG GTTTTATTCC ATATAATAAC

                PacI
               ~~~~~~~~~
35451  ATGATGTTAA TTAAGAATTC GGATCTGCGA CGCGAGGCTG GATGGCCTTC
       TACTACAATT AATTCTTAAG CCTAGACGCT GCGCTCCGAC CTACCGGAAG

35501  CCCATTATGA TTCTTCTCGC TTCCGGCGGC ATCGGGATGC CCGCGTTGCA
       GGGTAATACT AAGAAGAGCG AAGGCCGCCG TAGCCCTACG GGCGCAACGT

35551  GGCCATGCTG TCCAGGCAGG TAGATGACGA CCATCAGGGA CAGCTTCAAG
       CCGGTACGAC AGGTCCGTCC ATCTACTGCT GGTAGTCCCT GTCGAAGTTC
```

FIG.9A-42

35601 GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GGCGTTTTTC
CGGTCGTTTT CCGGTCCTTG GCATTTTTCC GGCGCAACGA CCGCAAAAAG

35651 CATAGGCTCC GCCCCCCTGA CGAGCATCAC AAAAATCGAC GCTCAAGTCA
GTATCCGAGG CGGGGGGACT GCTCGTAGTG TTTTTAGCTG CGAGTTCAGT

35701 GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG TTTCCCCCTG
CTCCACCGCT TTGGGCTGTC CTGATATTTC TATGGTCCGC AAAGGGGGAC

35751 GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA CCCTGCCGCT TACCGGATAC
CTTCGAGGGA GCACGCGAGA GGACAAGGCT GGGACGGCGA ATGGCCTATG

35801 CTGTCCGCCT TTCTCCCTTC GGGAAGCGTG GCGCTTTCTC ATAGCTCACG
GACAGGCGGA AAGAGGGAAG CCCTTCGCAC CGCGAAAGAG TATCGAGTGC

35851 CTGTAGGTAT CTCAGTTCGG TGTAGGTCGT TCGCTCCAAG CTGGGCTGTG
GACATCCATA GAGTCAAGCC ACATCCAGCA AGCGAGGTTC GACCCGACAC

35901 TGCACGAACC CCCCGTTCAG CCCGACCGCT GCGCCTTATC CGGTAACTAT
ACGTGCTTGG GGGGCAAGTC GGGCTGGCGA CGCGGAATAG GCCATTGATA

35951 CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC TGGCAGCAGC
GCAGAACTCA GGTTGGGCCA TTCTGTGCTG AATAGCGGTG ACCGTCGTCG

36001 CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT GCTACAGAGT
GTGACCATTG TCCTAATCGT CTCGCTCCAT ACATCCGCCA CGATGTCTCA

36051 TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGGAC AGTATTTGGT
AGAACTTCAC CACCGGATTG ATGCCGATGT GATCTTCCTG TCATAAACCA

36101 ATCTGCGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC
TAGACGCGAG ACGACTTCGG TCAATGGAAG CCTTTTTCTC AACCATCGAG

36151 TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTTT TTTGTTTGCA
AACTAGGCCG TTTGTTTGGT GGCGACCATC GCCACCAAAA AAACAAACGT

36201 AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA TCCTTTGATC
TCGTCGTCTA ATGCGCGTCT TTTTTTCCTA GAGTTCTTCT AGGAAACTAG

36251 TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT
AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTTGAGTG CAATTCCCTA

36301 TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAATC
AAACCAGTAC TCTAATAGTT TTTCCTAGAA GTGGATCTAG GAAAATTTAG

36351 AATCTAAAGT ATATATGAGT AAACTTGGTC TGACAGTTAC CAATGCTTAA
TTAGATTTCA TATATACTCA TTTGAACCAG ACTGTCAATG GTTACGAATT

36401 TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCGTTC ATCCATAGTT
AGTCACTCCG TGGATAGAGT CGCTAGACAG ATAAAGCAAG TAGGTATCAA

FIG.9A-43

```
36451  GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC
       CGGACTGAGG GGCAGCACAT CTATTGATGC TATGCCCTCC CGAATGGTAG

36501  TGGCCCCAGT GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG
       ACCGGGGTCA CGACGTTACT ATGGCGCTCT GGGTGCGAGT GGCCGAGGTC

36551  ATTTATCAGC AATAAACCAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT
       TAAATAGTCG TTATTTGGTC GGTCGGCCTT CCCGGCTCGC GTCTTCACCA

36601  CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAATTGTT GCCGGGAAGC
       GGACGTTGAA ATAGGCGGAG GTAGGTCAGA TAATTAACAA CGGCCCTTCG

36651  TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG
       ATCTCATTCA TCAAGCGGTC AATTATCAAA CGCGTTGCAA CAACGGTAAC

36701  CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC
       GATGTCCGTA GCACCACAGT GCGAGCAGCA AACCATACCG AAGTAAGTCG

36751  TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTTGTGCAA
       AGGCCAAGGG TTGCTAGTTC CGCTCAATGT ACTAGGGGGT ACAACACGTT

36801  AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG
       TTTTCGCCAA TCGAGGAAGC CAGGAGGCTA GCAACAGTCT TCATTCAACC

36851  CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT
       GGCGTCACAA TAGTGAGTAC CAATACCGTC GTGACGTATT AAGAGAATGA

36901  GTCATGCCAT CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA
       CAGTACGGTA GGCATTCTAC GAAAAGACAC TGACCACTCA TGAGTTGGTT

36951  GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT TGCCCGGCGT
       CAGTAAGACT CTTATCACAT ACGCCGCTGG CTCAACGAGA ACGGGCCGCA

37001  CAACACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC
       GTTGTGCCCT ATTATGGCGC GGTGTATCGT CTTGAAATTT TCACGAGTAG

37051  ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT
       TAACCTTTTG CAAGAAGCCC CGCTTTTGAG AGTTCCTAGA ATGGCGACAA

37101  GAGATCCAGT TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT
       CTCTAGGTCA AGCTACATTG GGTGAGCACG TGGGTTGACT AGAAGTCGTA

37151  CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT
       GAAAATGAAA GTGGTCGCAA AGACCCACTC GTTTTTGTCC TTCCGTTTTA

37201  GCCGCAAAAA AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT
       CGGCGTTTTT TCCCTTATTC CCGCTGTGCC TTTACAACTT ATGAGTATGA

37251  CTTCCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA
       GAAGGAAAAA GTTATAATAA CTTCGTAAAT AGTCCCAATA ACAGAGTACT
```

FIG.9A-44

37301 GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTTCCG
CGCCTATGTA TAAACTTACA TAAATCTTTT TATTTGTTTA TCCCCAAGGC

37351 CGCACATTTC CCCGAAAAGT GCCACCTGAC GTCTAAGAAA CCATTATTAT
GCGTGTAAAG GGGCTTTTCA CGGTGGACTG CAGATTCTTT GGTAATAATA

37401 CATGACATTA ACCTATAAAA ATAGGCGTAT CACGAGGCCC TTTCGTCTTC
GTACTGTAAT TGGATATTTT TATCCGCATA GTGCTCCGGG AAAGCAGAAG

37451 AAGAATTGGA TCCGAATTCT TAAT
TTCTTAACCT AGGCTTAAGA ATTA

FIG.9A-45

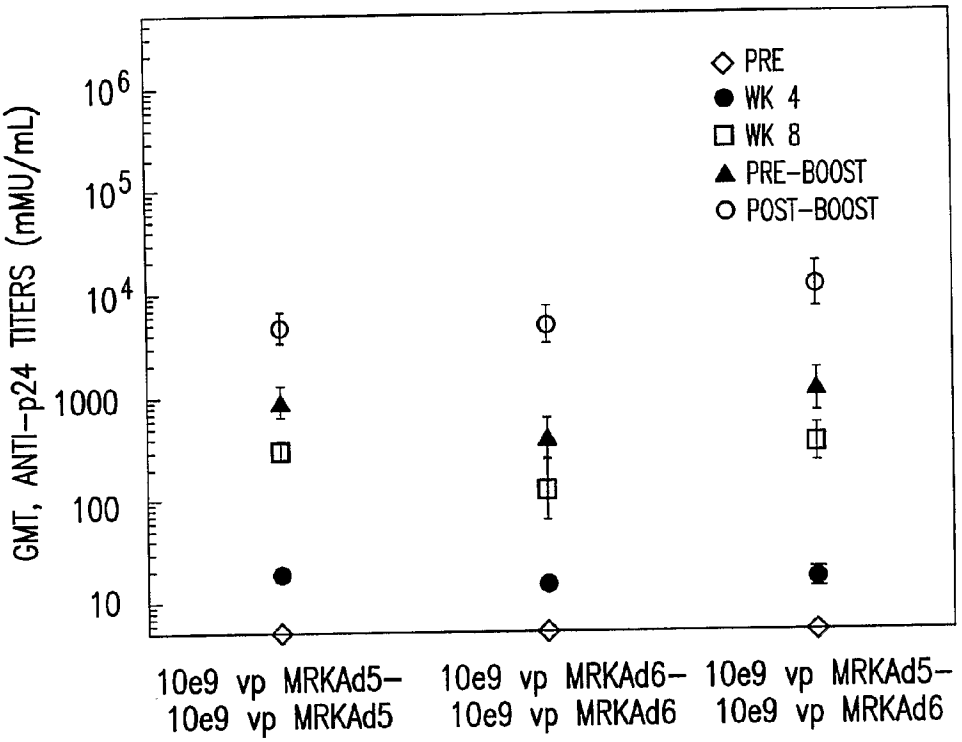

GMT, ANTI-p24 TITERS (mMU/mL)
10^6
10^5
10^4
1000
100
10
PRE
WK 4
WK 8
PRE-BOOST
POST-BOOST
10e9 vp MRKAd5-
10e9 vp MRKAd5
10e9 vp MRKAd6-
10e9 vp MRKAd6
10e9 vp MRKAd5-
10e9 vp MRKAd6

FIG.10

```
   1 CATCATCAAT AATATACCTT ATTTTGGATT GAAGCCAATA TGATAATGAG GGGGTGGAGT
  61 TTGTGACGTG GCGCGGGGCG TGGGAACGGG GCGGGTGACG TAGTAGTGTG GCGGAAGTGT
 121 GATGTTGTAA GTGTGGCGGA ACACATGTAA GCGCCGGATG TGGTAAAAGT GACGTTTTTG
 181 GTGTGCGCCG GTGTACACGG GAAGTGACAA TTTTCGCGCG GTTTTAGGCG GATGTTGTAG
 241 TAAATTTGGG CGTAACCAAG TAATATTTGG CCATTTTCGC GGGAAAACTG AATAAGAGGA
 301 AGTGAAATCT GAATAATTCT GTGTTACTCA TAGCGCGTAA TATTTGTCTA GGGCCGCGGG
 361 GACTTTGACC GTTTACGTGG AGACTCGCCC AGGTGTTTTT CTCAGGTGTT TTCCGCGTTC
 421 CGGGTCAAAG TTGGCGTTTT ATTATTATAG TCAGCTGACG CGCAGTGTAT TTATACCCGG
 481 TGAGTTCCTC AAGAGGCCAC TCTTGAGTGC CAGCGAGTAG AGTTTTCTCC TCCGAGCCGC
 541 TCCGACACCG GGACTGAAAA TGAGACATAT TATCTGCCAC GGAGGTGTTA TTACCGAAGA
 601 AATGGCCGCC AGTCTTTTGG ACCAGCTGAT CGAAGAGGTA CTGGCTGATA ATCTTCCACC
 661 TCCTAGCCAT TTTGAACCAC CTACCCTTCA CGAACTGTAT GATTTAGACG TGACGGCCCC
 721 CGAAGATCCC AACGAGGAGG CGGTTTCGCA GATTTTTCCC GAGTCTGTAA TGTTGGCGGT
 781 GCAGGAAGGG ATTGACTTAT TCACTTTTCC GCCGGCGCCC GGTTCTCCGG AGCCGCCTCA
 841 CCTTTCCCGG CAGCCCGAGC AGCCGGAGCA GAGAGCCTTG GGTCCGGTTT CTATGCCAAA
 901 CCTTGTGCCG GAGGTGATCG ATCTTACCTG CCACGAGGCT GGCTTTCCAC CCAGTGACGA
 961 CGAGGATGAA GAGGGTGAGG AGTTTGTGTT AGATTATGTG GAGCACCCCG GGCACGGTTG
1021 CAGGTCTTGT CATTATCACC GGAGGAATAC GGGGGACCCA GATATTATGT GTTCGCTTTG
1081 CTATATGAGG ACCTGTGGCA TGTTTGTCTA CAGTAAGTGA AAAATTATGG GCAGTGGGTG
1141 ATAGAGTGGT GGGTTTGGTG TGGTAATTTT TTTTTTAATT TTTACAGTTT TGTGGTTTAA
1201 AGAATTTTGT ATTGTGATTT TTTAAAAGGT CCTGTGTCTG AACCTGAGCC TGAGCCCGAG
1261 CCAGAACCGG AGCCTGCAAG ACCTACCCGG CGTCCTAAAT TGGTGCCTGC TATCCTGAGA
1321 CGCCCGACAT CACCTGTGTC TAGAGAATGC AATAGTAGTA CGGATAGCTG TGACTCCGGT
1381 CCTTCTAACA CACCTCCTGA GATACACCCG GTGGTCCCGC TGTGCCCCAT TAAACCAGTT
1441 GCCGTGAGAG TTGGTGGGCG TCGCCAGGCT GTGGAATGTA TCGAGGACTT GCTTAACGAG
1501 TCTGGGCAAC CTTTGGACTT GAGCTGTAAA CGCCCCAGGC CATAAGGTGT AAACCTGTGA
1561 TTGCGTGTGT GGTTAACGCC TTTGTTTGCT GAATGAGTTG ATGTAAGTTT AATAAAGGGT
1621 GAGATAATGT TTAACTTGCA TGGCGTGTTA AATGGGGCGG GGCTTAAAGG GTATATAATG
1681 CGCCGTGGGC TAATCTTGGT TACATCTGAC CTCATGGAGG CTTGGGAGTG TTTGGAAGAT
1741 TTTTCTGCTG TGCGTAACTT GCTGGAACAG AGCTCTAACA GTACCTCTTG GTTTTGGAGG
1801 TTTCTGTGGG GCTCCTCCCA GGCAAAGTTA GTCTGCAGAA TTAAGGAGGA TTACAAGTGG
1861 GAATTTGAAG AGCTTTTGAA ATCCTGTGGT GAGCTGTTTG ATTCTTTGAA TCTGGGTCAC
1921 CAGGCGCTTT TCCAAGAGAA GGTCATCAAG ACTTTGGATT TTTCCACACC GGGGCGCGCT
1981 GCGGCTGCTG TTGCTTTTTT GAGTTTTATA AAGGATAAAT GGAGCGAAGA AACCCATCTG
2041 AGCGGGGGGT ACCTGCTGGA TTTTCTGGCC ATGCATCTGT GGAGAGCGGT GGTGAGACAC
2101 AAGAATCGCC TGCTACTGTT GTCTTCCGTC CGCCCGGCAA TAATACCGAC GGAGGAGCAA
2161 CAGCAGGAGG AAGCCAGGCG GCGGCGGCGG CAGGAGCAGA GCCCATGGAA CCCGAGAGCC
2221 GGCCTGGACC CTCGGGAATG AATGTTGTAC AGGTGGCTGA ACTGTTTCCA GAACTGAGAC
2281 GCATTTTAAC CATTAACGAG GATGGGCAGG GGCTAAAGGG GGTAAAGAAG GAGCGGGGGG
2341 CTTCTGAGGC TACAGAGGAG GCTAGGAATC TAACTTTTAG CTTAATGACC AGACACCGTC
2401 CTGAGTGTGT TACTTTTCAG CAGATTAAGG ATAATTGCGC TAATGAGCTT GATCTGCTGG
2461 CGCAGAAGTA TTCCATAGAG CAGCTGACCA CTTACTGGCT GCAGCCAGGG GATGATTTTG
2521 AGGAGGCTAT TAGGGTATAT GCAAAGGTGG CACTTAGGCC AGATTGCAAG TACAAGATTA
2581 GCAAACTTGT AAATATCAGG AATTGTTGCT ACATTTCTGG GAACGGGGCC GAGGTGGAGA
2641 TAGATACGGA GGATAGGGTG GCCTTTAGAT GTAGCATGAT AAATATGTGG CCGGGGGTGC
```

FIG.11A-1

2701 TTGGCATGGA CGGGGTGGTT ATTATGAATG TGAGGTTTAC TGGTCCCAAT TTTAGCGGTA
2761 CGGTTTTCCT GGCCAATACC AATCTTATCC TACACGGTGT AAGCTTCTAT GGGTTTAACA
2821 ATACCTGTGT GGAAGCCTGG ACCGATGTAA GGGTTCGGGG CTGTGCCTTT TACTGCTGCT
2881 GGAAGGGGGT GGTGTGTCGC CCCAAAAGCA GGGCTTCAAT TAAGAAATGC CTGTTTGAAA
2941 GGTGTACCTT GGGTATCCTG TCTGAGGGTA ACTCCAGGGT GCGCCACAAT GTGGCCTCCG
3001 ACTGTGGTTG CTTTATGCTA GTGAAAAGCG TGGCTGTGAT TAAGCATAAC ATGGTGTGTG
3061 GCAACTGCGA GGACAGGGCC TCTCAGATGC TGACCTGCTC GGACGGCAAC TGTCACTTGC
3121 TGAAGACCAT TCACGTAGCC AGCCACTCTC GCAAGGCCTG GCCAGTGTTT GAGCACAACA
3181 TACTGACCCG CTGTTCCTTG CATTTGGGTA ACAGGAGGGG GGTGTTCCTA CCTTACCAAT
3241 GCAATTTGAG TCACACTAAG ATATTGCTTG AGCCCGAGAG CATGTCCAAG GTGAACCTGA
3301 ACGGGGTGTT TGACATGACC ATGAAGATCT GGAAGGTGCT GAGGTACGAT GAGACCCGCA
3361 CCAGGTGCAG ACCCTGCGAG TGTGGCGGTA AACATATTAG GAACCAGCCT GTGATGCTGG
3421 ATGTGACCGA GGAGCTGAGG CCCGATCACT TGGTGCTGGC CTGCACCCGC GCTGAGTTTG
3481 GCTCTAGCGA TGAAGATACA GATTGAGGTA CTGAAATGTG TGGGCGTGGC TTAAGGGTGG
3541 GAAAGAATAT ATAAGGTGGG GGTCTCATGT AGTTTTGTAT CTGTTTTGCA GCAGCCGCCG
3601 CCATGAGCGC CAACTCGTTT GATGGAAGCA TTGTGAGCTC ATATTTGACA ACGCGCATGC
3661 CCCCATGGGC CGGGGTGCGT CAGAATGTGA TGGGCTCCAG CATTGATGGT CGCCCCGTCC
3721 TGCCCGCAAA CTCTACTACC TTGACCTACG AGACCGTGTC TGGAACGCCG TTGGAGACTG
3781 CAGCCTCCGC CGCCGCTTCA GCCGCTGCAG CCACCGCCCG CGGGATTGTG ACTGACTTTG
3841 CTTTCCTGAG CCCGCTTGCA AGCAGTGCAG CTTCCCGTTC ATCCGCCCGC GATGACAAGT
3901 TGACGGCTCT TTTGGCACAA TTGGATTCTT TGACCCGGGA ACTTAATGTC GTTTCTCAGC
3961 AGCTGTTGGA TCTGCGCCAG CAGGTTTCTG CCCTGAAGGC TTCCTCCCCT CCCAATGCGG
4021 TTTAAAACAT AAATAAAAAC CAGACTCTGT TTGGATTTGG ATCAAGCAAG TGTCTTGCTG
4081 TCTTTATTTA GGGGTTTTGC GCGCGCGGTA GGCCCGGGAC CAGCGGTCTC GGTCGTTGAG
4141 GGTCCTGTGT ATTTTTTCCA GGACGTGGTA AAGGTGACTC TGGATGTTCA GATACATGGG
4201 CATAAGCCCG TCTCTGGGGT GGAGGTAGCA CCACTGCAGA GCTTCATGCT GCGGGGTGGT
4261 GTTGTAGATG ATCCAGTCGT AGCAGGAGCG CTGGGCGTGG TGCCTAAAAA TGTCTTTCAG
4321 TAGCAAGCTG ATTGCCAGGG GCAGGCCCTT GGTGTAAGTG TTTACAAAGC GGTTAAGCTG
4381 GGATGGGTGC ATACGTGGGG ATATGAGATG CATCTTGGAC TGTATTTTTA GGTTGGCTAT
4441 GTTCCCAGCC ATATCCCTCC GGGGATTCAT GTTGTGCAGA ACCACCAGCA CAGTGTATCC
4501 GGTGCACTTG GGAAATTTGT CATGTAGCTT AGAAGGAAAT GCGTGGAAGA ACTTGGAGAC
4561 GCCCTTGTGA CCTCCAAGAT TTTCCATGCA TTCGTCCATA ATGATGGCAA TGGGCCCACG
4621 GGCGGCGGCC TGGGCGAAGA TATTTCTGGG ATCACTAACG TCATAGTTGT GTTCCAGGAT
4681 GAGATCGTCA TAGGCCATTT TTACAAAGCG CGGGCGGAGG GTGCCAGACT GCGGTATAAT
4741 GGTTCCATCC GGCCCAGGGG CGTAGTTACC CTCACAGATT TGCATTTCCC ACGCTTTGAG
4801 TTCAGATGGG GGGATCATGT CTACCTGCGG GGCGATGAAG AAAACCGTTT CCGGGGTAGG
4861 GGAGATCAGC TGGGAAGAAA GCAGGTTCCT AAGCAGCTGC GACTTACCGC AGCCGGTGGG
4921 CCCGTAAATC ACACCTATTA CCGGCTGCAA CTGGTAGTTA AGAGAGCTGC AGCTGCCGTC
4981 ATCCCTGAGC AGGGGGGCCA CTTCGTTAAG CATGTCCCTG ACTTGCATGT TTTCCCTGAC
5041 CAAATCCGCC AGAAGGCGCT CGCCGCCCAG CGATAGCAGT TCTTGCAAGG AAGCAAAGTT
5101 TTTCAACGGT TTGAGGCCGT CCGCCGTAGG CATGCTTTTG AGCGTTTGAC CAAGCAGTTC
5161 CAGGCGGTCC CACAGCTCGG TCACGTGCTC TACGGCATCT CGATCCAGCA TATCTCCTCG
5221 TTTCGCGGGT TGGGGCGGCT TTCGCTGTAC GGCAGTAGTC GGTGCTCGTC CAGACGGGCC
5281 AGGGTCATGT CTTTCCACGG GCGCAGGGTC CTCGTCAGCG TAGTCTGGGT CACGGTGAAG
5341 GGGTGCGCTC CGGGTTGCGC GCTGGCCAGG GTGCGCTTGA GGCTGGTCCT GCTGGTGCTG

FIG.11A-2

5401 AAGCGCTGCC GGTCTTCGCC CTGCGCGTCG GCCAGGTAGC ATTTGACCAT GGTGTCATAG
5461 TCCAGCCCCT CCGCGGCGTG GCCCTTGGCG CGCAGCTTGC CCTTGGAGGA GGCGCCGCAC
5521 GAGGGGCAGT GCAGACTTTT AAGGGCGTAG AGCTTGGGCG CGAGAAATAC CGATTCCGGG
5581 GAGTAGGCAT CCGCGCCGCA GGCCCCGCAG ACGGTCTCGC ATTCCACGAG CCAGGTGAGC
5641 TCTGGCCGTT CGGGGTCAAA AACCAGGTTT CCCCCATGCT TTTTGATGCG TTTCTTACCT
5701 CTGGTTTCCA TGAGCCGGTG TCCACGCTCG GTGACGAAAA GGCTGTCCGT GTCCCCGTAT
5761 ACAGACTTGA GAGGCCTGTC CTCGAGCGGT GTTCCGCGGT CCTCCTCGTA TAGAAACTCG
5821 GACCACTCTG AGACGAAGGC TCGCGTCCAG GCCAGCACGA AGGAGGCTAA GTGGGAGGGG
5881 TAGCGGTCGT TGTCCACTAG GGGGTCCACT CGCTCCAGGG TGTGAAGACA CATGTCGCCC
5941 TCTTCGGCAT CAAGGAAGGT GATTGGTTTA TAGGTGTAGG CCACGTGACC GGGTGTTCCT
6001 GAAGGGGGGC TATAAAAGGG GGTGGGGGCG CGTTCGTCCT CACTCTCTTC CGCATCGCTG
6061 TCTGCGAGGG CCAGCTGTTG GGGTGAGTAC TCCCTCTCAA AAGCGGGCAT GACTTCTGCG
6121 CTAAGATTGT CAGTTTCCAA AAACGAGGAG GATTTGATAT TCACCTGGCC CGCGGTGATG
6181 CCTTTGAGGG TGGCCGCGTC CATCTGGTCA GAAAAGACAA TCTTTTTGTT GTCAAGCTTG
6241 GTGGCAAACG ACCCGTAGAG GGCGTTGGAC AGCAACTTGG CGATGGAGCG CAGGGTTTGG
6301 TTTTTGTCGC GATCGGCGCG CTCCTTGGCC GCGATGTTTA GCTGCACGTA TTCGCGCGCA
6361 ACGCACCGCC ATTCGGGAAA GACGGTGGTG CGCTCGTCGG GCACTAGGTG CACGCGCCAA
6421 CCGCGGTTGT GCAGGGTGAC AAGGTCAACG CTGGTGGCTA CCTCTCCGCG TAGGCGCTCG
6481 TTGGTCCAGC AGAGGCGGCC GCCCTTGCGC GAGCAGAATG GCGGTAGTGG GTCTAGCTGC
6541 GTCTCGTCCG GGGGGTCTGC GTCCACGGTA AAGACCCCGG GCAGCAGGCG CGCGTCGAAG
6601 TAGTCTATCT TGCATCCTTG CAAGTCTAGC GCCTGCTGCC ATGCGCGGGC GGCAAGCGCG
6661 CGCTCGTATG GGTTGAGTGG GGGACCCCAT GGCATGGGGT GGGTGAGCGC GGAGGCGTAC
6721 ATGCCGCAAA TGTCGTAAAC GTAGAGGGGC TCTCTGAGTA TTCCAAGATA TGTAGGGTAG
6781 CATCTTCCAC CGCGGATGCT GGCGCGCACG TAATCGTATA GTTCGTGCGA GGGAGCGAGG
6841 AGGTCGGGAC CGAGGTTGCT ACGGGCGGGC TGCTCTGCTC GGAAGACTAT CTGCCTGAAG
6901 ATGGCATGTG AGTTGGATGA TATGGTTGGA CGCTGGAAGA CGTTGAAGCT GGCGTCTGTG
6961 AGACCTACCG CGTCACGCAC GAAGGAGGCG TAGGAGTCGC GCAGCTTGTT GACCAGCTCG
7021 GCGGTGACCT GCACGTCTAG GGCGCAGTAG TCCAGGGTTT CCTTGATGAT GTCATACTTA
7081 TCCTGTCCCT TTTTTTTCCA CAGCTCGCGG TTGAGGACAA ACTCTTCGCG GTCTTTCCAG
7141 TACTCTTGGA TCGGAAACCC GTCGGCCTCC GAACGGTAAG AGCCTAGCAT GTAGAACTGG
7201 TTGACGGCCT GGTAGGCGCA GCATCCCTTT TCTACGGGTA GCGCGTATGC CTGCGCGGCC
7261 TTCCGGAGCG AGGTGTGGGT GAGCGCAAAG GTGTCCCTAA CCATGACTTT GAGGTACTGG
7321 TATTTGAAGT CAGTGTCGTC GCATCCGCCC TGCTCCCAGA GCAAAAAGTC CGTGCGCTTT
7381 TTGGAACGCG GGTTTGGCAG GGCGAAGGTG ACATCGTTGA AGAGTATCTT TCCCGCGCGA
7441 GGCATAAAGT TGCGTGTGAT GCGGAAGGGT CCCGGCACCT CGGAACGGTT GTTAATTACC
7501 TGGGCGGCGA GCACGATCTC GTCAAAGCCG TTGATGTTGT GGCCCACAAT GTAAAGTTCC
7561 AAGAAGCGCG GGATGCCCTT GATGGAAGGC AATTTTTTAA GTTCCTCGTA GGTGAGCTCT
7621 TCAGGGGAGC TGAGCCCGTG CTCTGAAAGG GCCCAGTCTG CAAGATGAGG GTTGGAAGCG
7681 ACGAATGAGC TCCACAGGTC ACGGGCCATT AGCATTTGCA GGTGGTCGCG AAAGGTCCTA
7741 AACTGGCGAC CTATGGCCAT TTTTTCTGGG GTGATGCAGT AGAAGGTAAG CGGGTCTTGT
7801 TCCCAGCGGT CCCATCCAAG GTCCGCGGCT AGGTCTCGCG CGGCGGTCAC TAGAGGCTCA
7861 TCTCCGCCGA ACTTCATGAC CAGCATGAAG GGCACGAGCT GCTTCCCAAA GGCCCCCATC
7921 CAAGTATAGG TCTCTACATC GTAGGTGACA AAGAGACGCT CGGTGCGAGG ATGCGAGCCG
7981 ATCGGGAAGA ACTGGATCTC CCGCCACCAG TTGGAGGAGT GGCTGTTGAT GTGGTGAAAG
8041 TAGAAGTCCC TGCGACGGGC CGAACACTCG TGCTGGCTTT TGTAAAAACG TGCGCAGTAC

FIG.11A-3

8101 TGGCAGCGGT GCACGGGCTG TACATCCTGC ACGAGGTTGA CCTGACGACC GCGCACAAGG
8161 AAGCAGAGTG GGAATTTGAG CCCCTCGCCT GGCGGGTTTG GCTGGTGGTC TTCTACTTCG
8221 GCTGCTTGTC CTTGACCGTC TGGCTGCTCG AGGGGAGTTA CGGTGGATCG GACCACCACG
8281 CCGCGCGAGC CCAAAGTCCA GATGTCCGCG CGCGGCGGTC GGAGCTTGAT GACAACATCG
8341 CGCAGATGGG AGCTGTCCAT GGTCTGGAGC TCCCGCGGCG TCAGGTCAGG CGGGAGCTCC
8401 TGCAGGTTTA CCTCGCATAG CCGGGTCAGG GCGCGGGCTA GGTCCAGGTG ATACCTGATT
8461 TCCAGGGGCT GGTTGGTGGC GGCGTCGATG GCTTGCAAGA GGCCGCATCC CCGCGGCGCG
8521 ACTACGGTAC CGCGCGGCGG GCGGTGGGCC GCGGGGGTGT CCTTGGATGA TGCATCTAAA
8581 AGCGGTGACG CGGGCGGGCC CCCGGAGGTA GGGGGGGCTC GGGACCCGCC GGGAGAGGGG
8641 GCAGGGGCAC GTCGGCGCCG CGCGCGGGCA GGAGCTGGTG CTGCGCGCGG AGGTTGCTGG
8701 CGAACGCGAC GACGCGGCGG TTGATCTCCT GAATCTGGCG CCTCTGCGTG AAGACGACGG
8761 GCCCGGTGAG CTTGAACCTG AAAGAGAGTT CGACAGAATC AATTTCGGTG TCGTTGACGG
8821 CGGCCTGGCG CAAAATCTCC TGCACGTCTC CTGAGTTGTC TTGATAGGCG ATCTCGGCCA
8881 TGAACTGCTC GATCTCTTCC TCCTGGAGAT CTCCGCGTCC GGCTCGCTCC ACGGTGGCGG
8941 CGAGGTCGTT GGAGATGCGG GCCATGAGCT GCGAGAAGGC GTTGAGGCCT CCCTCGTTCC
9001 AGACGCGGCT GTAGACCACG CCCCCTTCGG CATCGCGGGC GCGCATGACC ACCTGCGCGA
9061 GATTGAGCTC CACGTGCCGG GCGAAGACGG CGTAGTTTCG CAGGCGCTGA AAGAGGTAGT
9121 TGAGGGTGGT GGCGGTGTGT TCTGCCACGA AGAAGTACAT AACCCAGCGC CGCAACGTGG
9181 ATTCGTTGAT ATCCCCCAAG GCCTCAAGGC GCTCCATGGC CTCGTAGAAG TCCACGGCGA
9241 AGTTGAAAAA CTGGGAGTTG CGCGCCGACA CGGTTAACTC CTCCTCCAGA AGACGGATGA
9301 GCTCGGCGAC AGTGTCGCGC ACCTCGCGCT CAAAGGCTAC AGGGGCCTCT TCTTCTTCTT
9361 CAATCTCCTC TTCCATAAGG GCCTCCCCTT CTTCTTCTTC TGGCGGCGGT GGGGGAGGGG
9421 GGACACGGCG GCGACGACGG CGCACCGGGA GGCGGTCGAC AAAGCGCTCG ATCATCTCCC
9481 CGCGGCGACG GCGCATGGTC TCGGTGACGG CGCGGCCGTT CTCGCGGGGG CGCAGTTGGA
9541 AGACGCCGCC CGTCATGTCC CGGTTATGGG TTGGCGGGGG GCTGCCGTGC GGCAGGGATA
9601 CGGCGCTAAC GATGCATCTC AACAATTGTT GTGTAGGTAC TCCGCCACCG AGGGACCTGA
9661 GCGAGTCCGC ATCGACCGGA TCGGAAAACC TCTCGAGAAA GGCGTCTAAC CAGTCACAGT
9721 CGCAAGGTAG GCTGAGCACC GTGGCGGGCG GCAGCGGGCG GCGGTCGGGG TTGTTTCTGG
9781 CGGAGGTGCT GCTGATGATG TAATTAAAGT AGGCGGTCTT GAGACGGCGG ATGGTCGACA
9841 GAAGCACCAT GTCCTTGGGT CCGGCCTGCT GAATGCGCAG GCGGTCGGCC ATGCCCCAGG
9901 CTTCGTTTTG ACATCGGCGC AGGTCTTTGT AGTAGTCTTG CATGAGCCTT TCTACCGGCA
9961 CTTCTTCTTC TCCTTCCTCT TGTCCTGCAT CTCTTGCATC TATCGCTGCG GCGGCGGCGG
10021 AGTTTGGCCG TAGGTGGCGC CCTCTTCCTC CCATGCGTGT GACCCCGAAG CCCCTCATCG
10081 GCTGAAGCAG GGCCAGGTCG GCGACAACGC GCTCGGCTAA TATGGCCTGC TGCACCTGCG
10141 TGAGGGTAGA CTGGAAGTCG TCCATGTCCA CAAAGCGGTG GTATGCGCCC GTGTTGATGG
10201 TGTAAGTGCA GTTGGCCATA ACGGACCAGT TAACGGTCTG GTGACCCGGC TGCGAGAGCT
10261 CGGTGTACCT GAGACGCGAG TAAGCCCTTG AGTCAAAGAC GTAGTCGTTG CAAGTCCGCA
10321 CCAGGTACTG GTATCCCACC AAAAAGTGCG GCGGCGGCTG GCGGTAGAGG GGCCAGCGTA
10381 GGGTGGCCGG GGCTCCGGGG GCGAGGTCTT CCAACATAAG GCGATGATAT CCGTAGATGT
10441 ACCTGGACAT CCAGGTGATG CCGGCGGCGG TGGTGGAGGC GCGCGGAAAG TCACGGACGC
10501 GGTTCCAGAT GTTGCGCAGC GGCAAAAAGT GCTCCATGGT CGGGACGCTC TGGCCGGTCA
10561 GGCGCGCGCA GTCGTTGACG CTCTAGACCG TGCAAAAGGA GAGCCTGTAA GCGGGCACTC
10621 TTCCGTGGTC TGGTGGATAA ATTCGCAAGG GTATCATGGC GGACGACCGG GGTTCGAACC
10681 CCGGATCCGG CCGTCCGCCG TGATCCATGC GGTTACCGCC CGCGTGTCGA ACCCAGGTGT
10741 GCGACGTCAG ACAACGGGGG AGCGCTCCTT TTGGCTTCCT TCCAGGCGCG GCGGATGCTG

FIG.11A-4

10801 CGCTAGCTTT TTTGGCCACT GGCCGCGCGC GGCGTAAGCG GTTAGGCTGG AAAGCGAAAG
10861 CATTAAGTGG CTCGCTCCCT GTAGCCGGAG GGTTATTTTC CAAGGGTTGA GTCGCGGGAC
10921 CCCCGGTTCG AGTCTCGGGC CGGCCGGACT GCGGCGAACG GGGGTTTGCC TCCCCGTCAT
10981 GCAAGACCCC GCTTGCAAAT TCCTCCGGAA ACAGGGACGA GCCCCTTTTT TGCTTTTCCC
11041 AGATGCATCC GGTGCTGCGG CAGATGCGCC CCCCTCCTCA GCAGCGGCAA GAGCAAGAGC
11101 AGCGGCAGAC ATGCAGGGCA CCCTCCCCTT CTCCTACCGC GTCAGGAGGG GCAACATCCG
11161 CGGCTGACGC GGCGGCAGAT GGTGATTACG AACCCCCGCG GCGCCGGACC CGGCACTACT
11221 TGGACTTGGA GGAGGGCGAG GGCCTGGCGC GGCTAGGAGC GCCCTCTCCT GAGCGACACC
11281 CAAGGGTGCA GCTGAAGCGT GACACGCGCG AGGCGTACGT GCCGCGGCAG AACCTGTTTC
11341 GCGACCGCGA GGGAGAGGAG CCCGAGGAGA TGCGGGATCG AAAGTTCCAT GCAGGGCGCG
11401 AGTTGCGGCA TGGCCTGAAC CGCGAGCGGT TGCTGCGCGA GGAGGACTTT GAGCCCGACG
11461 CGCGGACCGG GATTAGTCCC GCGCGCGCAC ACGTGGCGGC CGCCGACCTG GTAACCGCGT
11521 ACGAGCAGAC GGTGAACCAG GAGATTAACT TTCAAAAAAG CTTTAACAAC CACGTGCGCA
11581 CGCTTGTGGC GCGCGAGGAG GTGGCTATAG GACTGATGCA TCTGTGGGAC TTTGTAAGCG
11641 CGCTGGAGCA AAACCCAAAT AGCAAGCCGC TCATGGCGCA GCTGTTCCTT ATAGTGCAGC
11701 ACAGCAGGGA CAACGAGGCA TTCAGGGATG CGCTGCTAAA CATAGTAGAG CCCGAGGGCC
11761 GCTGGCTGCT CGATTTGATA AACATTCTGC AGAGCATAGT GGTGCAGGAG CGCAGCTTGA
11821 GCCTGGCTGA CAAGGTGGCC GCCATTAACT ATTCCATGCT CAGTCTGGGC AAGTTTTACG
11881 CCCGCAAGAT ATACCATACC CCTTACGTTC CCATAGACAA GGAGGTAAAG ATCGAGGGGT
11941 TCTACATGCG CATGGCGCTG AAGGTGCTTA CCTTGAGCGA CGACCTGGGC GTTTATCGCA
12001 ACGAGCGCAT CCACAAGGCC GTGAGCGTGA GCCGGCGGCG CGAGCTCAGC GACCGCGAGC
12061 TGATGCACAG CCTGCAAAGG GCCCTGGCTG GCACGGGCAG CGGCGATAGA GAGGCCGAGT
12121 CCTACTTTGA CGCGGGCGCT GACCTGCGCT GGGCCCCAAG CCGACGCGCC CTGGAGGCAG
12181 CTGGGGCCGG ACCTGGGCTG GCGGTGGCAC CCGCGCGCGC TGGCAACGTC GGCGGCGTGG
12241 AGGAATATGA CGAGGACGAT GAGTACGAGC CAGAGGACGG CGAGTACTAA GCGGTGATGT
12301 TTCTGATCAG ATGATGCAAG ACGCAACGGA CCCGGCGGTG CGGGCGGCGC TGCAGAGCCA
12361 GCCGTCCGGC CTTAACTCCA CGGACGACTG GCGCCAGGTC ATGGACCGCA TCATGTCGCT
12421 GACTGCGCGC AACCCTGACG CGTTCCGGCA GCAGCCGCAG GCCAACCGGC TCTCCGCAAT
12481 TCTGGAAGCG GTGGTCCCGG CGCGCGCAAA CCCCACGCAC GAGAAGGTGC TGGCGATCGT
12541 AAACGCGCTG GCCGAAAACA GGGCCATCCG GCCCGATGAG GCCGGCCTGG TCTACGACGC
12601 GCTGCTTCAG CGCGTGGCTC GTTACAACAG CAGCAACGTG CAGACCAACC TGGACCGGCT
12661 GGTGGGGGAT GTGCGCGAGG CCGTGGCGCA GCGTGAGCGC GCGCAGCAGC AGGGCAACCT
12721 GGGCTCCATG GTTGCACTAA ACGCCTTCCT GAGTACACAG CCCGCCAACG TGCCGCGGGG
12781 ACAGGAGGAC TACACCAACT TTGTGAGCGC ACTGCGGCTA ATGGTGACTG AGACACCGCA
12841 AAGTGAGGTG TATCAGTCCG GGCCAGACTA TTTTTTCCAG ACCAGTAGAC AAGGCCTGCA
12901 GACCGTAAAC CTGAGCCAGG CTTTCAAGAA CTTGCAGGGG CTGTGGGGGG TGCGGGCTCC
12961 CACAGGCGAC CGCGCGACCG TGTCTAGCTT GCTGACGCCC AACTCGCGCC TGTTGCTGCT
13021 GCTAATAGCG CCCTTCACGG ACAGTGGCAG CGTGTCCCGG GACACATACC TAGGTCACTT
13081 GCTGACACTG TACCGCGAGG CCATAGGTCA GGCGCATGTG GACGAGCATA CTTTCCAGGA
13141 GATTACAAGT GTTAGCCGCG CGCTGGGGCA GGAGGACACG GGCAGCCTGG AGGCAACCCT
13201 GAACTACCTG CTGACCAACC GGCGGCAAAA AATCCCCTCG TTGCACAGTT TAAACAGCGA
13261 GGAGGAGCGC ATTTTGCGCT ATGTGCAGCA GAGCGTGAGC CTTAACCTGA TGCGCGACGG
13321 GGTAACGCCC AGCGTGGCGC TGGACATGAC CGCGCGCAAC ATGGAACCGG GCATGTATGC
13381 CTCAAACCGG CCGTTTATCA ATCGCCTAAT GGACTACTTG CATCGCGCGG CCGCCGTGAA
13441 CCCCGAGTAT TTCACCAATG CCATCTTGAA CCCGCACTGG CTACCGCCCC CTGGTTTCTA

FIG.11A-5

13501 CACCGGGGGA TTCGAGGTGC CCGAGGGTAA CGATGGATTC CTCTGGGACG ACATAGACGA
13561 CAGCGTGTTT TCCCCGCAAC CGCAGACCCT GCTAGAGTTG CAACAACGCG AGCAGGCAGA
13621 GGCGGCGCTG CGAAAGGAAA GCTTCCGCAG GCCAAGCAGC TTGTCCGATC TAGGCGCTGC
13681 GGCCCCGCGG TCAGATGCTA GTAGCCCATT TCCAAGCTTG ATAGGGTCTC TTACCAGCAC
13741 TCGCACCACC CGCCCGCGCC TGCTGGGCGA GGAGGAGTAC CTAAACAACT CGCTGCTGCA
13801 GCCGCAGCGC GAAAAGAACC TGCCTCCGGC GTTTCCCAAC AACGGGATAG AGAGCCTAGT
13861 GGACAAGATG AGTAGATGGA AGACGTATGC GCAGGAGCAC AGGGATGTGC CCGGCCCGCG
13921 CCCGCCCACC CGTCGTCAAA GGCACGACCG TCAGCGGGGT CTGGTGTGGG AGGACGATGA
13981 CTCGGCAGAC GACAGCAGCG TCTTGGATTT GGGAGGGAGT GGCAACCCGT TTGCACACCT
14041 TCGCCCCAGG CTGGGGAGAA TGTTTTAAAA AAAGCATGAT GCAAAATAAA AAACTCACCA
14101 AGGCCATGGC ACCGAGCGTT GGTTTTCTTG TATTCCCCTT AGTATGCGGC GCGCGGCGAT
14161 GTATGAGGAA GGTCCTCCTC CCTCCTACGA GAGCGTGGTG AGCGCGGCGC CAGTGGCGGC
14221 GGCGCTGGGT TCACCCTTCG ATGCTCCCCT GGACCCGCCG TTCGTGCCTC CGCGGTACCT
14281 GCGGCCTACC GGGGGGAGAA ACAGCATCCG TTACTCTGAG TTGGCACCCC TATTCGACAC
14341 CACCCGTGTG TACCTTGTGG ACAACAAGTC AACGGATGTG GCATCCCTGA ACTACCAGAA
14401 CGACCACAGC AACTTTCTAA CCACGGTCAT TCAAAACAAT GACTACAGCC CGGGGGAGGC
14461 AAGCACACAG ACCATCAATC TTGACGACCG GTCGCACTGG GGCGGCGACC TGAAAACCAT
14521 CCTGCATACC AACATGCCAA ATGTGAACGA GTTCATGTTT ACCAATAAGT TTAAGGCGCG
14581 GGTGATGGTG TCGCGCTCGC TTACTAAGGA CAAACAGGTG GAGCTGAAAT ACGAGTGGGT
14641 GGAGTTCACG CTGCCCGAGG GCAACTACTC CGAGACCATG ACCATAGACC TTATGAACAA
14701 CGCGATCGTG GAGCACTACT TGAAAGTGGG CAGGCAGAAC GGGGTTCTGG AAAGCGACAT
14761 CGGGGTAAAG TTTGACACCC GCAACTTCAG ACTGGGGTTT GACCCAGTCA CTGGTCTTGT
14821 CATGCCTGGG GTATATACAA ACGAAGCCTT CCATCCAGAC ATCATTTTGC TGCCAGGATG
14881 CGGGGTGGAC TTCACCCACA GCCGCCTGAG CAACTTGTTG GGCATCCGCA AGCGGCAACC
14941 CTTCCAGGAG GGCTTTAGGA TCACCTACGA TGACCTGGAG GGTGGTAACA TTCCCGCACT
15001 GTTGGATGTG GACGCCTACC AGGCAAGCTT GAAAGATGAC ACCGAACAGG GCGGGGGTGG
15061 CGCAGGCGGC GGCAACAACA GTGGCAGCGG CGCGGAAGAG AACTCCAACG CGGCAGCTGC
15121 GGCAATGCAG CCGGTGGAGG ACATGAACGA TCATGCCATT CGCGGCGACA CCTTTGCCAC
15181 ACGGGCGGAG GAGAAGCGCG CTGAGGCCGA GGCAGCGGCC GAAGCTGCCG CCCCCGCTGC
15241 GGAGGCTGCA CAACCCGAGG TCGAGAAGCC TCAGAAGAAA CCGGTGATTA AACCCCTGAC
15301 AGAGGACAGC AAGAAACGCA GTTACAACCT AATAAGCAAT GACAGCACCT TCACCCAGTA
15361 CCGCAGCTGG TACCTTGCAT ACAACTACGG CGACCCTCAG GCCGGGATCC GCTCATGGAC
15421 CCTGCTTTGC ACTCCTGACG TAACCTGCGG CTCGGAGCAG GTATACTGGT CGTTGCCCGA
15481 CATGATGCAA GACCCCGTGA CCTTCCGCTC CACGCGCCAG ATCAGCAACT TTCCGGTGGT
15541 GGGCGCCGAG CTGTTGCCCG TGCACTCCAA GAGCTTCTAC AACGACCAGG CCGTCTACTC
15601 CCAGCTCATC CGCCAGTTTA CCTCTCTGAC CCACGTGTTC AATCGCTTTC CCGAGAACCA
15661 GATTTTGGCG CGCCCGCCAG CCCCCACCAT CACCACCGTC AGTGAAAACG TTCCTGCTCT
15721 CACAGATCAC GGGACGCTAC CGCTGCGCAA CAGCATCGGA GGAGTCCAGC GAGTGACCAT
15781 TACTGACGCC AGACGCCGCA CCTGCCCCTA CGTTTACAAG GCCCTGGGCA TAGTCTCGCC
15841 GCGCGTCCTA TCGAGCCGCA CTTTTTGAGC AAGCATGTCC ATCCTTATAT CGCCCAGCAA
15901 TAACACAGGC TGGGGCCTGC GCTTCCCAAG CAAGATGTTT GGCGGGGCCA AGAAGCGCTC
15961 CGACCAACAC CCAGTGCGCG TGCGCGGGCA CTACCGCGCG CCCTGGGGCG CGCACAAACG
16021 CGGCCGCACT GGGCGCACCA CCGTCGATGA CGCCATCGAC GCGGTGGTGG AGGAGGCGCG
16081 CAACTACACG CCCACGCCGC CGCCAGTGTC CACCGTGGAC GCGGCCATTC AGACCGTGGT
16141 GCGCGGAGCC CGGCGCTACG CTAAAATGAA GAGACGGCGG AGGCGCGTAG CACGTCGCCA

FIG.11A-6

16201 CCGCCGCCGA CCCGGCACTG CCGCCCAACG CGCGGCGGCG GCCCTGCTTA ACCGCGCACG
16261 TCGCACCGGC CGACGGGCGG CCATGCGAGC CGCTCGAAGG CTGGCCGCGG GTATTGTCAC
16321 TGTGCCCCCC AGGTCCAGGC GACGAGCGGC CGCCGCAGCA GCCGCGGCCA TTAGTGCTAT
16381 GACTCAGGGT CGCAGGGGCA ACGTGTACTG GGTGCGCGAC TCGGTTAGCG GCCTGCGCGT
16441 GCCCGTGCGC ACCCGCCCCC CGCGCAACTA GATTGCAATA AAAAACTACT TAGACTCGTA
16501 CTGTTGTATG TATCCAGCGG CGGCGGCGCG CATCGAAGCT ATGTCCAAGC GCAAAATCAA
16561 AGAAGAGATG CTCCAGGTCA TCGCGCCGGA GATCTATGGC CCCCCGAAGA AGGAAGAGCA
16621 GGATTACAAG CCCCGAAAGC TAAAGCGGGT CAAAAAGAAA AAGAAAGATG ATGATGATGA
16681 TGAACTTGAC GACGAGGTGG AACTGTTGCA CGCGACCGCG CCCAGGCGAC GGGTACAGTG
16741 GAAAGGTCGA CGCGTAAGAC GTGTTTTGCG ACCCGGCACC ACCGTAGTCT TTACGCCCGG
16801 TGAGCGCTCC ACCCGCACCT ACAAGCGCGT GTATGATGAG GTGTACGGCG ACGAGGACCT
16861 GCTTGAGCAG GCCAACGAGC GCCTCGGGGA GTTTGCCTAC GGAAAGCGGC ATAAGGACAT
16921 GCTGGCGTTG CCGCTGGACG AGGGCAACCC AACACCTAGC CTAAAGCCCG TGACACTGCA
16981 GCAGGTGCTG CCCGCGCTTG CACCGTCCGA AGAAAAGCGC GGCCTAAAGC GCGAGTCTGG
17041 TGACTTGGCA CCCACCGTGC AGCTGATGGT ACCCAAGCGT CAGCGACTGG AAGATGTCTT
17101 GGAAAAHATG ACCGTGGAGC CTGGGCTGGA GCCCGAGGTC CGCGTGCGGC CAATCAAGCA
17161 GGTGGCACCG GGACTGGGCG TGCAGACCGT GGACGTTCAG ATACCCACCA CCAGTAGCAC
17221 TAGTATTGCC ACTGCCACAG AGGGCATGGA GACACAAACG TCCCCGGTTG CCTCGGCGGT
17281 GGCAGATGCC GCGGTGCAGG CGGCCGCTGC GGCCGCGTCC AAGACCTCTA CGGAGGTGCA
17341 AACGGACCCG TGGATGTTTC GTGTTTCAGC CCCCCGGCGT CCGCGCCGTT CAAGGAAGTA
17401 CGGCGCCGCC AGCGCGCTAC TGCCCGAATA TGCCCTACAT CCTTCCATCG CGCCTACCCC
17461 CGGCTATCGT GGCTACACCT ACCGCCCCAG AAGACGAGCA ACTACCCGAC GCCGAACCAC
17521 CACTGGAACC CGCCGCCGCC GTCGCCGTCG CCAGCCCGTG CTGGCCCCGA TTTCCGTGCG
17581 CAGGGTGGCT CGCGAAGGAG GCAGGACCCT GGTGCTGCCA ACAGCGCGCT ACCACCCCAG
17641 CATCGTTTAA AAGCCGGTCT TTGTGGTTCT TGCAGATATG GCCCTCACCT GCCGCCTCCG
17701 TTTCCCGGTG CCGGGATTCC GAGGAAGAAT GCACCGTAGG AGGGGCATGG CCGGCCACGG
17761 CCTGACGGGC GGCATGCGTC GTGCGCACCA CCGGCGGCGG CGCGCGTCGC ACCGTCGCAT
17821 GCGCGGCGGT ATCCTGCCCC TCCTTATTCC ACTGATCGCC GCGGCGATTG GCGCCGTGCC
17881 CGGAATTGCA TCCGTGGCCT TGCAGGCGCA GAGACACTGA TTAAAAACAA GTTACATGTG
17941 GAAAAATCAA AATAAAAGTC TGGACTCTCA CGCTCGCTTG GTCCTGTAAC TATTTTGTAG
18001 AATGGAAGAC ATCAACTTTG CGTCACTGGC CCCGCGACAC GGCTCGCGCC CGTTCATGGG
18061 AAACTGGCAA GATATCGGCA CCAGCAATAT GAGCGGTGGC GCCTTCAGCT GGGGCTCGCT
18121 GTGGAGCGGC ATTAAAAATT TCGGTTCCGC CGTTAAGAAC TATGGCAGCA AAGCCTGGAA
18181 CAGCAGCACA GGCCAGATGC TGAGGGACAA GTTGAAAGAG CAAAATTTCC AACAAAAGGT
18241 GGTAGATGGC CTGGCCTCTG GCATTAGCGG GGTGGTGGAC CTGGCCAACC AGGCAGTGCA
18301 AAATAAGATT AACAGTAAGC TTGATCCCCG CCCTCCCGTA GAGGAGCCTC CACCGGCCGT
18361 GGAGACAGTG TCTCCAGAGG GGCGTGGCGA AAAGCGTCCG CGACCCGACA GGGAAGAAAC
18421 TCTGGTGACG CAAATAGACG AGCCTCCCTC GTACGAGGAG GCACTAAAGC AAGGCCTGCC
18481 CACCACCCGT CCCATCGCGC CCATGGCTAC CGGAGTGCTG GGCCAGCACA CACCCGTAAC
18541 GCTGGACCTG CCTCCCCCCG CCGACACCCA GCAGAAACCT GTGCTGCCAG GCCCGTCCGC
18601 CGTTGTTGTA ACCCGTCCTA GCCGCGCGTC CCTGCGCCGC GCCGCCAGCG GTCCGCGATC
18661 GTTGCGGCCC GTAGCCAGTG GCAACTGGCA AAGCACACTG AACAGCATCG TGGGTTTGGG
18721 GGTGCAATCC CTGAAGCGCC GACGATGCTT CTGATAGCTA ACGTGTCGTA TGTGTGTCAT
18781 GTATGCGTCC ATGTCGCCGC CAGAGGAGCT GCTGAGCCGC CGCGCGCCCG CTTTCCAAGA
18841 TGGCTACCCC TTCGATGATG CCGCAGTGGT CTTACATGCA CATCTCGGGC CAGGACGCCT

FIG.11A-7

18901 CGGAGTACCT GAGCCCCGGG CTGGTGCAGT TCGCCCGCGC CACCGAGACG TACTTCAGCC
18961 TGAATAACAA GTTTAGAAAC CCCACGGTGG CGCCTACGCA CGACGTGACC ACAGACCGGT
19021 CTCAGCGTTT GACGCTGCGG TTCATCCCCG TGGACCGCGA GGATACTGCG TACTCGTACA
19081 AGGCGCGGTT CACCCTAGCT GTGGGTGATA ACCGTGTGCT AGACATGGCT TCCACGTACT
19141 TTGACATCCG CGGCGTGCTG GACAGGGGCC CTACTTTTAA GCCCTACTCT GGCACTGCCT
19201 ACAACGCACT GGCCCCCAAG GGTGCCCCCA ACTCGTGCGA GTGGGAACAA AATGAAACTG
19261 CACAAGTGGA TGCTCAAGAA CTTGACGAAG AGGAGAATGA AGCCAATGAA GCTCAGGCGC
19321 GAGAACAGGA ACAAGCTAAG AAAACCCATG TATATGCCCA GGCTCCACTG TCCGGAATAA
19381 AAATAACTAA AGAAGGTCTA CAAATAGGAA CTGCCGACGC CACAGTAGCA GGTGCCGGCA
19441 AAGAAATTTT CGCAGACAAA ACTTTTCAAC CTGAACCACA AGTAGGAGAA TCTCAATGGA
19501 ACGAAGCGGA TGCCACAGCA GCTGGTGGAA GGGTTCTTAA AAAGACAACT CCCATGAAAC
19561 CCTGCTATGG CTCATACGCT AGACCCACCA ATTCCAACGG CGGACAGGGC GTTATGGTTG
19621 AACAAAATGG TAAATTGGAA AGTCAAGTCG AAATGCAATT TTTTTCCACA TCCACAAATG
19681 CCACAAATGA AGTTAACAAT ATACAACCAA CAGTTGTATT GTACAGCGAA GATGTAAACA
19741 TGGAAACTCC AGATACTCAT CTTTCTTATA AACCTAAAAT GGGGGATAAA AATGCCAAAG
19801 TCATGCTTGG ACAACAAGCA ATGCCAAACA GACCAAATTA CATTGCTTTT AGAGACAATT
19861 TTATTGGTCT CATGTATTAC AACAGCACAG GTAACATGGG TGTCCTTGCT GGTCAGGCAT
19921 CGCAGTTGAA CGCTGTTGTA GATTTGCAAG ACAGAAACAC AGAGCTGTCC TACCAGCTTT
19981 TGCTTGATTC AATTGGCGAC AGAACAAGAT ACTTTTCAAT GTGGAATCAA GCTGTTGACA
20041 GCTATGATCC AGATGTCAGA ATTATTGAGA ACCATGGAAC TGAGGATGAG TTGCCAAATT
20101 ATTGCTTTCC TCTTGGTGGA ATTGGGATTA CTGACACTTT TCAAGCTGTT AAAACAACTG
20161 CTGCTAACGG GGACCAAGGC AATACTACCT GGCAAAAAGA TTCAACATTT GCAGAACGCA
20221 ATGAAATAGG GGTGGGAAAT AACTTTGCCA TGGAAATTAA CCTGAATGCC AACCTATGGA
20281 GAAATTTCCT TTACTCCAAT ATTGCGCTGT ACCTGCCAGA CAAGCTAAAA TACAACCCCA
20341 CCAATGTGGA AATATCTGAC AACCCCAACA CCTACGACTA CATGAACAAG CGAGTGGTGG
20401 CTCCTGGGCT TGTAGACTGC TACATTAACC TTGGGGCGCG CTGGTCTCTG GACTACATGG
20461 ACAACGTTAA TCCCTTTAAC CACCACCGCA ATGCGGGCCT GCGTTACCGC TCCATGTTGT
20521 TGGGAAACGG CCGCTACGTG CCCTTTCACA TTCAGGTGCC CCAAAAGTTT TTTGCCATTA
20581 AAAACCTCCT CCTCCTGCCA GGCTCATACA CATATGAATG GAACTTCAGG AAGGATGTTA
20641 ACATGGTTCT GCAGAGCTCT CTGGGAAACG ACCTTAGAGT TGACGGGGCT AGCATTAAGT
20701 TTGACAGCAT TTGTCTTTAC GCCACCTTCT TCCCCATGGC CCACAACACG GCCTCCACGC
20761 TGGAAGCCAT GCTCAGAAAT GACACCAACG ACCAGTCCTT TAATGACTAC CTTTCCGCCG
20821 CCAACATGCT ATATCCCATA CCCGCCAACG CCACCAACGT GCCCATCTCC ATCCCATCGC
20881 GCAACTGGGC AGCATTTCGC GGTTGGGCCT TCACACGCTT GAAGACAAAG GAAACCCCTT
20941 CCCTGGGATC AGGCTACGAC CCTTACTACA CCTACTCTGG CTCCATACCA TACCTTGACG
21001 GAACCTTCTA TCTTAATCAC ACCTTTAAGA AGGTGGCCAT TACTTTTGAC TCTTCTGTTA
21061 GCTGGCCGGG CAACGACCGC CTGCTTACTC CCAATGAGTT TGAGATTAAG CGCTCAGTTG
21121 ACGGGGAGGG CTATAACGTA GCTCAGTGCA ACATGACAAA GGACTGGTTC CTAGTGCAGA
21181 TGTTGGCCAA CTACAATATT GGCTACCAGG GCTTCTACAT TCCAGAAAGC TACAAAGACC
21241 GCATGTACTC GTTCTTCAGA AACTTCCAGC CCATGAGCCG GCAAGTGGTG GACGATACTA
21301 AATACAAAGA TTATCAGCAG GTTGGAATTA TCCACCAGCA TAACAACTCA GGCTTCGTAG
21361 GCTACCTCGC TCCCACCATG CGCGAGGGAC AAGCTTACCC CGCTAATGTT CCCTACCCAC
21421 TAATAGGCAA AACCGCGGTT GATAGTATTA CCCAGAAAAA GTTTCTTTGC GACCGCACCC
21481 TGTGGCGCAT CCCCTTCTCC AGTAACTTTA TGTCCATGGG TGCGCTCACA GACCTGGGCC
21541 AAAACCTTCT CTACGCAAAC TCCGCCCACG CGCTAGACAT GACCTTTGAG GTGGATCCCA

FIG.11A-8

21601 TGGACGAGCC CACCCTTCTT TATGTTTTGT TTGAAGTCTT TGACGTGGTC CGTGTGCACC
21661 AGCCGCACCG CGGCGTCATC GAGACCGTGT ACCTGCGCAC GCCCTTCTCG GCCGGCAACG
21721 CCACAACATA AAGAAGCAAG CAACATCAAC AACAGCTGCC GCCATGGGCT CCAGTGAGCA
21781 GGAACTGAAA GCCATTGTCA AAGATCTTGG TTGTGGGCCA TATTTTTTGG GCACCTATGA
21841 CAAGCGCTTC CCAGGCTTTG TTTCCCCACA CAAGCTCGCC TGCGCCATAG TTAACACGGC
21901 CGGTCGCGAG ACTGGGGGCG TACACTGGAT GGCCTTTGCC TGGAACCCGC GCTCAAAAAC
21961 ATGCTACCTC TTTGAGCCCT TTGGCTTTTC TGACCAACGT CTCAAGCAGG TTTACCAGTT
22021 TGAGTACGAG TCACTCCTGC GCCGTAGCGC CATTGCCTCT TCCCCCGACC GCTGTATAAC
22081 GCTGGAAAAG TCCACCCAAA GCGTGCAGGG GCCCAACTCG GCCGCCTGTG GCCTATTCTG
22141 CTGCATGTTT CTCCACGCCT TTGCCAACTG GCCCCAAACT CCCATGGATC ACAACCCCAC
22201 CATGAACCTT ATTACCGGGG TACCCAACTC CATGCTTAAC AGTCCCCAGG TACAGCCCAC
22261 CCTGCGCCGC AACCAGGAAC AGCTCTACAG CTTCCTGGAG CGCCACTCGC CCTACTTCCG
22321 CAGCCACAGT GCGCAAATTA GGAGCGCCAC TTCTTTTTGT CACTTGAAAA ACATGTAAAA
22381 ATAATGTACT AGGAGACACT TTCAATAAAG GCAAATGTTT TTATTTGTAC ACTCTCGGGT
22441 GATTATTTAC CCCCACCCTT GCCGTCTGCG CCGTTTAAAA ATCAAAGGGG TTCTGCCGCG
22501 CATCGCTATG CGCCACTGGC AGGGACACGT TGCGATACTG GTGTTTAGTG CTCCACTTAA
22561 ACTCAGGCAC AACCATCCGC GGCAGCTCGG TGAAGTTTTC ACTCCACAGG CTGCGCACCA
22621 TCACCAACGC GTTTAGCAGG TCGGGCGCCG ATATCTTGAA GTCGCAGTTG GGGCCTCCGC
22681 CCTGCGCGCG CGAGTTGCGA TACACAGGGT TACAGCACTG GAACACTATC AGCGCCGGGT
22741 GGTGCACGCT GGCCAGCACG CTCTTGTCGG AGATCAGATC CGCGTCCAGG TCCTCCGCGT
22801 TGCTCAGGGC GAACGGAGTC AACTTTGGTA GCTGCCTTCC CAAAAAGGGT GCATGCCCAG
22861 GCTTTGAGTT GCACTCGCAC CGTAGTGGCA TCAGAAGGTG ACCGTGCCCA GTCTGGGCGT
22921 TAGGATACAG CGCCTGCATG AAAGCCTTGA TCTGCTTAAA AGCCACCTGA GCCTTTGCGC
22981 CTTCAGAGAA GAACATGCCG CAAGACTTGC CGGAAAACTG ATTGGCCGGA CAGGCCGCGT
23041 CATGCACGCA GCACCTTGCG TCGGTGTTGG AGATCTGCAC CACATTTCGG CCCCACCGGT
23101 TCTTCACGAT CTTGGCCTTG CTAGACTGCT CCTTCAGCGC GCGCTGCCCG TTTTCGCTCG
23161 TCACATCCAT TTCAATCACG TGCTCCTTAT TTATCATAAT GCTCCCGTGT AGACACTTAA
23221 GCTCGCCTTC GATCTCAGCG CAGCGGTGCA GCCACAACGC GCAGCCCGTG GGCTCGTGGT
23281 GCTTGTAGGT TACCTCTGCA AACGACTGCA GGTACGCCTG CAGGAATCGC CCCATCATCG
23341 TCACAAAGGT CTTGTTGCTG GTGAAGGTCA GCTGCAACCC GCGGTGCTCC TCGTTTAGCC
23401 AGGTCTTGCA TACGGCCGCC AGAGCTTCCA CTTGGTCAGG CAGTAGCTTG AAGTTTGCCT
23461 TTAGATCGTT ATCCACGTGG TACTTGTCCA TCAACGCGCG CGCAGCCTCC ATGCCCTTCT
23521 CCCACGCAGA CACGATCGGC AGGCTCAGCG GGTTTATCAC CGTGCTTTCA CTTTCCGCTT
23581 CACTGGACTC TTCCTTTTCC TCTTGCATCC GCATACCCCG CGCCACTGGG TCGTCTTCAT
23641 TCAGCCGCCG CACCGTGCGC TTACCTCCCT TGCCGTGCTT GATTAGCACC GGTGGGTTGC
23701 TGAAACCCAC CATTTGTAGC GCCACATCTT CTCTTTCTTC CTCGCTGTCC ACGATCACCT
23761 CTGGGGATGG CGGGCGCTCG GGCTTGGGAG AGGGGCGCTT CTTTTTCTTT TTGGACGCAA
23821 TGGCCAAATC CGCCGTCGAG GTCGATGGCC GCGGGCTGGG TGTGCGCGGC ACCAGCGCAT
23881 CTTGTGACGA GTCTTCTTCG TCCTCGGACT CGAGACGCCG CCTCAGCCGC TTTTTTGGGG
23941 GCGCGCGGGG AGGCGGCGGC GACGGCGACG GGGACGAGAC GTCCTCCATG GTTGGTGGAC
24001 GTCGCGCCGC ACCGCGTCCG CGCTCGGGGG TGGTTTCGCG CTGCTCCTCT TCCCGACTGG
24061 CCATTTCCTT CTCCTATAGG CAGAAAAAGA TCATGGAGTC AGTCGAGAAG GAGGACAGCC
24121 TAACCGCCCC CTTTGAGTTC GCCACCACCG CCTCCACCGA TGCCGCCAAC GCGCCTACCA
24181 CCTTCCCCGT CGAGGCACCC CCGCTTGAGG AGGAGGAAGT GATTATCGAG CAGGACCCAG
24241 GTTTTGTAAG CGAAGACGAC GAAGATCGCT CAGTACCAAC AGAGGATAAA AAGCAAGACC

FIG.11A-9

24301 AGGACGACGC AGAGGCAAAC GAGGAACAAG TCGGGCGGGG GGACCAAAGG CATGGCGACT
24361 ACCTAGATGT GGGAGACGAC GTGCTGTTGA AGCATCTGCA GCGCCAGTGC GCCATTATCT
24421 GCGACGCGTT GCAAGAGCGC AGCGATGTGC CCCTCGCCAT AGCGGATGTC AGCCTTGCCT
24481 ACGAACGCCA CCTGTTCTCA CCGCGCGTAC CCCCCAAACG CCAAGAAAAC GGCACATGCG
24541 AGCCCAACCC GCGCCTCAAC TTCTACCCCG TATTTGCCGT GCCAGAGGTG CTTGCCACCT
24601 ATCACATCTT TTTCCAAAAC TGCAAGATAC CCCTATCCTG CCGTGCCAAC CGCAGCCGAG
24661 CGGACAAGCA GCTGGCCTTG CGGCAGGGCG CTGTCATACC TGATATCGCC TCGCTCGACG
24721 AAGTGCCAAA AATCTTTGAG GGTCTTGGAC GCGACGAGAA GCGCGCGGCA AACGCTCTGC
24781 AACAAGAAAA CAGCGAAAAT GAAAGTCACT GTGGAGTGCT GGTGGAACTT GAGGGTGACA
24841 ACGCGCGCCT AGCCGTGCTG AAACGCAGCA TCGAGGTCAC CCACTTTGCC TACCCGGCAC
24901 TTAACCTACC CCCCAAGGTT ATGAGCACAG TCATGAGCGA GCTGATCGTG CGCCGTGCAC
24961 GACCCCTGGA GAGGGATGCA AACTTGCAAG AACAAACCGA GGAGGGCCTA CCCGCAGTTG
25021 GCGATGAGCA GCTGGCGCGC TGGCTTGAGA CGCGCGAGCC TGCCGACTTG GAGGAGCGAC
25081 GCAAGCTAAT GATGGCCGCA GTGCTTGTTA CCGTGGAGCT TGAGTGCATG CAGCGGTTCT
25141 TTGCTGACCC GGAGATGCAG CGCAAGCTAG AGGAAACGTT GCACTACACC TTTCGCCAGG
25201 GCTACGTGCG CCAGGCCTGC AAAATTTCCA ACGTGGAGCT CTGCAACCTG GTCTCCTACC
25261 TTGGAATTTT GCACGAAAAC CGCCTTGGGC AAAACGTGCT TCATTCCACG CTCAAGGGCG
25321 AGGCGCGCCG CGACTACGTC CGCGACTGCG TTTACTTATT TCTGTGCTAC ACCTGGCAAA
25381 CGGCCATGGG CGTGTGGCAG CAGTGCCTGG AGGAGCGCAA CCTGAAGGAG CTGCAGAAGC
25441 TGCTAAAGCA AAACTTGAAG GACCTATGGA CGGCCTTCAA CGAGCGCTCC GTGGCCGCGC
25501 ACCTGGCGGA CATTATCTTC CCCGAACGCC TGCTTAAAAC CCTGCAACAG GGTCTGCCAG
25561 ACTTCACCAG TCAAAGCATG TTGCAAAACT TTAGGAACTT TATCCTAGAG CGTTCAGGAA
25621 TTCTGCCCGC CACCTGCTGT GCGCTTCCTA GCGACTTTGT GCCCATTAAG TACCGTGAAT
25681 GCCCTCCGCC GCTTTGGGGT CACTGCTACC TTCTGCAGCT AGCCAACTAC CTTGCCTACC
25741 ACTCCGACAT CATGGAAGAC GTGAGCGGTG ACGGCCTACT GGAGTGTCAC TGTCGCTGCA
25801 ACCTATGCAC CCCGCACCGC TCCCTGGTCT GCAATTCACA ACTGCTTAGC GAAAGTCAAA
25861 TTATCGGTAC CTTTGAGCTG CAGGGTCCCT CGCCTGACGA AAAGTCCGCG GCTCCGGGGT
25921 TGAAACTCAC TCCGGGGCTG TGGACGTCGG CTTACCTTCG CAAATTTGTA CCTGAGGACT
25981 ACCACGCCCA CGAGATTAGG TTCTACGAAG ACCAATCCCG CCCGCCAAAT GCGGAGCTTA
26041 CCGCCTGCGT CATTACCCAG GGCCACATCC TTGGCCAATT GCAAGCCATT AACAAAGCCC
26101 GCCAAGAGTT TCTGCTACGA AAGGGACGGG GGGTTTACTT GGACCCCCAG TCCGGCGAGG
26161 AGCTCAACCC AATCCCCCCG CCGCCGCAGC CCTATCAGCA GCCGCGGGCC CTTGCTTCCC
26221 AGGATGGCAC CCAAAAAGAA GCTGCAGCTG CCGCCGCCGC CACCCACGGA CGAGGAGGAA
26281 TACTGGGACA GTCAGGCAGA GGAGGTTTTG GACGAGGAGG AGGAGATGAT GGAAGACTGG
26341 GACAGCCTAG ACGAGGAAGC TTCCGAGGCC GAAGAGGTGT CAGACGAAAC ACCGTCACCC
26401 TCGGTCGCAT TCCCCTCGCC GGCGCCCCAG AAATCGGCAA CCGTTCCCAG CATTGCTACA
26461 ACCTCCGCTC CTCAGGCGCC GCCGGCACTG CCCGTTCGCC GACCCAACCG TAGATGGGAC
26521 ACCACTGGAA CCAGGGCCGG TAAGTCTAAG CAGCCGCCGC CGTTAGCCCA AGAGCAACAA
26581 CAGCGCCAAG GCTACCGCTC GTGGCGCGTG CACAAGAACG CCATAGTTGC TTGCTTGCAA
26641 GACTGTGGGG GCAACATCTC CTTCGCCCGC CGCTTTCTTC TCTACCATCA CGGCGTGGCC
26701 TTCCCCCGTA ACATCCTGCA TTACTACCGT CATCTCTACA GCCCCTACTG CACCGGCGGC
26761 AGCGGCAGCA ACAGCAGCGG CCACGCAGAA GCAAAGGCGA CCGGATAGCA AGACTCTGAC
26821 AAAGCCCAAG AAATCCACAG CGGCGGCAGC AGCAGGAGGA GGAGCACTGC GTCTGGCGCC
26881 CAACGAACCC GTATCGACCC GCGAGCTTAG AAACAGGATT TTTCCCACTC TGTATGCTAT
26941 ATTTCAACAG AGCAGGGGCC AAGAACAAGA GCTGAAAATA AAAAACAGGT CTCTGCGCTC

FIG.11A-10

27001 CCTCACCCGC AGCTGCCTGT ATCACAAAAG CGAAGATCAG CTTCGGCGCA CGCTGGAAGA
27061 CGCGGAGGCT CTCTTCAGCA AATACTGCGC GCTGACTCTT AAGGACTAGT TTCGCGCCCT
27121 TTCTCAAATT TAAGCGCGAA AACTACGTCA TCTCCAGCGG CCACACCCGG CGCCAGCACC
27181 TGTCGTCAGC GCCATTATGA GCAAGGAAAT TCCCACGCCC TACATGTGGA GTTACCAGCC
27241 ACAAATGGGA CTTGCGGCTG GAGCTGCCCA AGACTACTCA ACCCGAATAA ACTACATGAG
27301 CGCGGGACCC CACATGATAT CCCGGGTCAA CGGAATCCGC GCCCACCGAA ACCGAATTCT
27361 CCTCGAACAG GCGGCTATTA CCACCACACC TCGTAATAAC CTTAATCCCC GTAGTTGGCC
27421 CGCTGCCCTG GTGTACCAGG AAAGTCCCGC TCCCACCACT GTGGTACTTC CCAGAGACGC
27481 CCAGGCCGAA GTTCAGATGA CTAACTCAGG GGCGCAGCTT GCGGGCGGCT TTCGTCACAG
27541 GGTGCGGTCG CCCGGGCAGG GTATAACTCA CCTGAAAATC AGAGGGCGAG GTATTCAGCT
27601 CAACGACGAG TCGGTGAGCT CCTCTCTTGG TCTCCGTCCG GACGGGACAT TTCAGATCGG
27661 CGGCGCTGGC CGCTCTTCAT TTACGCCCCG TCAGGCGATC CTAACTCTGC AGACCTCGTC
27721 CTCGGAGCCG CGCTCCGGAG GCATTGGAAC TCTACAATTT ATTGAGGAGT TCGTGCCTTC
27781 GGTTTACTTC AACCCCTTTT CTGGACCTCC CGGCCACTAC CCGGACCAGT TTATTCCCAA
27841 CTTTGACGCG GTAAAAGACT CGGCGGACGG CTACGACTGA ATGACCAGTG GAGAGGCAGA
27901 GCAACTGCGC CTGACACACC TCGACCACTG CCGCCGCCAC AAGTGCTTTG CCCGCGGCTC
27961 CGGTGAGTTT TGTTACTTTG AATTGCCCGA AGAGCATATC GAGGGCCCGG CGCACGGCGT
28021 CCGGCTCACC ACCCAGGTAG AGCTTACACG TAGCCTGATT CGGGAGTTTA CCAAGCGCCC
28081 CCTGCTAGTG GAGCGGGAGC GGGGTCCCTG TGTTCTGACC GTGGTTTGCA ACTGTCCTAA
28141 CCCTGGATTA CATCAAGATC TTTGTTGTCA TCTCTGTGCT GAGTATAATA AATACAGAAA
28201 TTAGAATCTA CTGGGGCTCC TGTCGCCATC CTGTGAACGC CACCGTTTTT ACCCACCCAA
28261 AGCAGACCAA AGCAAACCTC ACCTCCGGTT TGCACAAGCG GGCCAATAAG TACCTTACCT
28321 GGTACTTTAA CGGCTCTTCA TTTGTAATTT ACAACAGTTT CCAGCGAGAC GAAGTAAGTT
28381 TGCCACACAA CCTTCTCGGC TTCAACTACA CCGTCAAGAA AAACACCACC ACCACCCTCC
28441 TCACCTGCCG GGAACGTACG AGTGCGTCAC CGGTTGCTGC GCCCACACCT ACAGCCTGAG
28501 CGTAACCAGA CATTACTCCC ATTTTCCCAA AACAGGAGGT GAGCTCAACT CCCGGAACTC
28561 AGGTCAAAAA AGCATTTTGC GGGGTGCTGG GATTTTTTAA TTAAGTATAT GAGCAATTCA
28621 AGTAACTCTA CAAGCTTGTC TAATTTTTCT GGAATTGGGG TCGGGGTTAT CCTTACTCTT
28681 GTAATTCTGT TTATTCTTAT ACTAGCACTT CTGTGCCTTA GGGTTGCCGC CTGCTGCACG
28741 CACGTTTGTA CCTATTGTCA GCTTTTTAAA CGCTGGGGGC GACATCCAAG ATGAGGTACA
28801 TGATTTTAGG CTTGCTCGCC CTTGCGGCAG TCTGCAGCGC TGCCAAAAAG GTTGAGTTTA
28861 AGGAACCAGC TTGCAATGTT ACATTTAAAT CAGAAGCTAA TGAATGCACT ACTCTTATAA
28921 AATGCACCAC AGAACATGAA AAGCTTATTA TTCGCCACAA AGACAAAATT GGCAAGTATG
28981 CTGTATATGC TATTTGGCAG CCAGGTGACA CTAACGACTA TAATGTCACA GTCTTCCAAG
29041 GTGAAAATCG TAAAACTTTT ATGTATAAAT TTCCATTTTA TGAAATGTGC GATATTACCA
29101 TGTACATGAG CAAACAGTAC AAGTTGTGGC CCCCACAAAA GTGTTTAGAG AACACTGGCA
29161 CCTTTTGTTC CACCGCTCTG CTTATTACAG CGCTTGCTTT GGTATGTACC TTACTTTATC
29221 TCAAATACAA AAGCAGACGC AGTTTTATTG ATGAAAAGAA AATGCCTTGA TTTTCCGCTT
29281 GCTTGTATTC CCCTGGACAA TTTACTCTAT GTGGGATATG CGCCAGGCGG GAAAGATTAT
29341 ACCCACAACC TTCAAATCAA ACTTTCCTGG ACGTTAGCGC CTGACTTCTG CCAGCGCCTG
29401 CACTGCAAAT TTGATCAAAC CCAGCTTCAG CTTGCCTGCT CCAGAGATGA CCGGCTCAAC
29461 CATCGCGCCC ACAACGGACT ATCGCAACAC CACTGCTACC GGACTAAAAT CTGCCCTAAA
29521 TTTACCCCAA GTTCATGCCT TTGTCAATGA CTGGGCGAGC TTGGGCATGT GGTGGTTTTC
29581 CATAGCGCTT ATGTTTGTTT GCCTTATTAT TATGTGGCTT ATTTGTTGCC TAAAGCGCAG
29641 ACGCGCCAGA CCCCCCATCT ATAGGCCTAT CATTGTGCTC AACCCACACA ATGAAAAAAT

FIG.11A-11

29701 TCATAGATTG GACGGTCTCA AACCATGTTC TCTTCTTTTA CAGTATGATT AAATGAGACA
29761 TGATTCCTCG AGTCCTTATA TTATTGACCC TTGTTGCGCT TTTCTGTGCG TGCTCTACAT
29821 TGGCTGCGGT CGCTCACATC GAAGTAGATT GCATCCCACC TTTCACAGTT TACCTGCTTT
29881 ACGGATTTGT CACCCTTATC CTCATCTGCA GCCTCGTCAC TGTAGTCATC GCCTTCATTC
29941 AGTTCATTGA CTGGATTTGT GTGCGCATTG CGTACCTTAG GCACCATCCG CAATACAGAG
30001 ACAGGACTAT AGCTGATCTT CTCAGAATTC TTTAATTATG AAACGGATTG TCACTTTTGT
30061 TTTGCTGATT TTCTGCGCCC TACCTGTGCT TTGCTCCCAA ACCTCAGCGC CTCCCAAAAG
30121 ACATATTTCC TGCAGATTCA CTCAAATATG GAACATTCCC AGCTGCTACA ACAAACAGAG
30181 CGATTTGTCA GAAGCCTGGT TATACGCCAT CATCTCTGTC ATGGTTTTTT GCAGTACCAT
30241 TTTTGCCCTA GCCATATACC CATACCTTGA CATTGGTTGG AATGCCATAG ATGCCATGAA
30301 CCACCCTACT TTCCCAGCGC CCAATGTCAT ACCACTGCAA CAGGTTATTG CCCCAATCAA
30361 TCAGCCTCGC CCCCCTTCTC CCACCCCCAC TGAGATTAGC TACTTTAATT TGACAGGTGG
30421 AGATGACTGA ATCTCTAGAT CTAGAATTGG ATGGAATTAA CACCGAACAG CGCCTACTAG
30481 AAAGGCGCAA GGCGGCGTCC GAGCGAGAAC GCCTAAAACA AGAAGTTGAA GACATGGTTA
30541 ACCTGCACCA GTGTAAAAGA GGTATCTTTT GTGTGGTCAA GCAGGCCAAA CTTACCTACG
30601 AAAAAACCAC TACCGGCAAC CGCCTTAGCT ACAAGCTACC CACCCAGCGC CAAAAACTGG
30661 TGCTTATGGT GGGAGAAAAA CCTATCACCG TCACCCAGCA CTCGGCAGAA ACAGAAGGCT
30721 GCCTGCACTT CCCCTATCAG GGTCCAGAGG ACCTCTGCAC TCTTATTAAA ACCATGTGTG
30781 GCATTAGAGA TCTTATTCCA TTCAACTAAC AATAAACACA CAATAAATTA CTTACTTAAA
30841 ATCAGTCAGC AAATCTTTGT CCAGCTTATT CAGCATCACC TCCTTTCCCT CCTCCCAACT
30901 CTGGTATTTC AGCAGCCTTT TAGCTGCGAA CTTTCTCCAA AGTCTAAATG GGATGTCAAA
30961 TTCCTCATGT TCTTGTCCCT CCGCACCCAC TATCTTCATA TTGTTGCAGA TGAAACGCGC
31021 CAGACCGTCT GAAGACACCT TCAACCCTGT GTACCCATAT GACACGGAAA CCGGCCCTCC
31081 AACTGTGCCT TTCCTTACCC CTCCCTTTGT GTCGCCAAAT GGGTTCCAAG AAAGTCCCCC
31141 CGGAGTGCTT TCTTTGCGTC TTTCAGAACC TTTGGTTACC TCACACGGCA TGCTTGCGCT
31201 AAAAATGGGC AGCGGCCTGT CCCTGGATCA GGCAGGCAAC CTTACATCAA ATACAATCAC
31261 TGTTTCTCAA CCGCTAAAAA AAACAAAGTC CAATATAACT TTGGAAACAT CCGCGCCCCT
31321 TACAGTCAGC TCAGGCGCCC TAACCATGGC CACAACTTCG CCTTTGGTGG TCTCTGACAA
31381 CACTCTTACC ATGCAATCAC AAGCACCGCT AACCGTGCAA GACTCAAAAC TTAGCATTGC
31441 TACCAAAGAG CCACTTACAG TGTTAGATGG AAAACTGGCC CTGCAGACAT CAGCCCCCCT
31501 CTCTGCCACT GATAACAACG CCCTCACTAT CACTGCCTCA CCTCCTCTTA CTACTGCAAA
31561 TGGTAGTCTG GCTGTTACCA TGGAAAACCC ACTTTACAAC AACAATGGAA AACTTGGGCT
31621 CAAAATTGGC GGTCCTTTGC AAGTGGCCAC CGACTCACAT GCACTAACAC TAGGTACTGG
31681 TCAGGGGGTT GCAGTTCATA ACAATTTGCT ACATACAAAA GTTACAGGCG CAATAGGGTT
31741 TGATACATCT GGCAACATGG AACTTAAAAC TGGAGATGGC CTCTATGTGG ATAGCGCCGG
31801 TCCTAACCAA AAACTACATA TTAATCTAAA TACCACAAAA GGCCTTGCTT TTGACAACAC
31861 CGCAATAACA ATTAACGCTG GAAAAGGGTT GGAATTTGAA ACAGACTCCT CAAACGGAAA
31921 TCCCATAAAA ACAAAAATTG GATCAGGCAT ACAATATAAT ACCAATGGAG CTATGGTTGC
31981 AAAACTTGGA ACAGGCCTCA GTTTTGACAG CTCCGGAGCC ATAACAATGG GCAGCATAAA
32041 CAATGACAGA CTTACTCTTT GGACAACACC AGACCCATCC CCAAATTGCA GAATTGCTTC
32101 AGATAAAGAC TGCAAGCTAA CTCTGGCGCT AACAAAATGT GGCAGTCAAA TTTTGGGCAC
32161 TGTTTCAGCT TTGGCAGTAT CAGGTAATAT GGCCTCCATC AATGGAACTC TAAGCAGTGT
32221 AAACTTGGTT CTTAGATTTG ATGACAACGG AGTGCTTATG TCAAATTCAT CACTGGACAA
32281 ACAGTATTGG AACTTTAGAA ACGGGGACTC CACTAACGGT CAACCATACA CTTATGCTGT
32341 TGGGTTTATG CCAAACCTAA AAGCTTACCC AAAAACTCAA AGTAAAACTG CAAAAAGTAA

FIG.11A-12

32401 TATTGTTAGC CAGGTGTATC TTAATGGTGA CAAGTCTAAA CCATTGCATT TTACTATTAC
32461 GCTAAATGGA ACAGATGAAA CCAACCAAGT AAGCAAATAC TCAATATCAT TCAGTTGGTC
32521 CTGGAACAGT GGACAATACA CTAATGACAA ATTTGCCACC AATTCCTATA CCTTCTCCTA
32581 CATTGCCCAG GAATAAAGAA TCGTGAACCT GTTGCATGTT ATGTTTCAAC GTGTTTATTT
32641 TTCAATTGCA GAAAATTTCA AGTCATTTTT CATTCAGTAG TATAGCCCCA CCACCACATA
32701 GCTTATACTA ATCACCGTAC CTTAATCAAA CTCACAGAAC CCTAGTATTC AACCTGCCAC
32761 CTCCCTCCCA ACACACAGAG TACACAGTCC TTTCTCCCCG GCTGGCCTTA AACAGCATCA
32821 TATCATGGGT AACAGACATA TTCTTAGGTG TTATATTCCA CACGGTCTCC TGTCGAGCCA
32881 AACGCTCATC AGTGATGTTA ATAAACTCCC CGGGCAGCTC GCTTAAGTTC ATGTCGCTGT
32941 CCAGCTGCTG AGCCACAGGC TGCTGTCCAA CTTGCGGTTG CTCAACGGGC GGCGAAGGAG
33001 AAGTCCACGC CTACATGGGG GTAGAGTCAT AATCGTGCAT CAGGATAGGG CGGTGGTGCT
33061 GCAGCAGCGC GCGAATAAAC TGCTGCCGCC GCCGCTCCGT CCTGCAGGAA TACAACATGG
33121 CAGTGGTCTC CTCAGCGATG ATTCGCACCG CCCGCAGCAT AAGGCGCCTT GTCCTCCGGG
33181 CACAGCAGCG CACCCTGATC TCACTTAAGT CAGCACAGTA ACTGCAGCAC AGTACCACAA
33241 TATTGTTTAA AATCCCACAG TGCAAGGCGC TGTATCCAAA GCTCATGGCG GGGACCACAG
33301 AACCCACGTG GCCATCATAC CACAAGCGCA GGTAGATTAA GTGGCGACCC CTCATAAACA
33361 CGCTGGACAT AAACATTACC TCTTTTGGCA TGTTGTAATT CACCACCTCC CGGTACCATA
33421 TAAACCTCTG ATTAAACATG GCGCCATCCA CCACCATCCT AAACCAGCTG GCCAAAACCT
33481 GCCCGCCGGC TATGCACTGC AGGGAACCGG GACTGGAACA ATGACAGTGG AGAGCCCAGG
33541 ACTCGTAACC ATGGATCATC ATGCTCGTCA TGATATCAAT GTTGGCACAA CACAGGCACA
33601 CGTGCATACA CTTCCTCAGG ATTACAAGCT CCTCCCGCGT CAGAACCATA TCCCAGGGAA
33661 CAACCCATTC CTGAATCAGC GTAAATCCCA CACTGCAGGG AAGACCTCGC ACGTAACTCA
33721 CGTTGTGCAT TGTCAAAGTG TTACATTCGG GCAGCAGCGG ATGATCCTCC AGTATGGTAG
33781 CGCGTGTCTC TGTCTCAAAA GGAGGTAGGC GATCCCTACT GTACGGAGTG CGCCGAGACA
33841 ACCGAGATCG TGTTGGTCGT AGTGTCATGC CAAATGGAAC GCCGGACGTA GTCATATTTC
33901 CTGAAGCAAA ACCAGGTGCG GGCGTGACAA ACAGATCTGC GTCTCCGGTC TCGTCGCTTA
33961 GCTCGCTCTG TGTAGTAGTT GTAGTATATC CACTCTCTCA AAGCATCCAG GCGCCCCCTG
34021 GCTTCGGGTT CTATGTAAAC TCCTTCATGC GCCGCTGCCC TGATAACATC CACCACCGCA
34081 GAATAAGCCA CACCCAGCCA ACCTACACAT TCGTTCTGCG AGTCACACAC GGGAGGAGCG
34141 GGAAGAGCTG GAAGAACCAT GTTTTTTTTT TTTATTCCAA AAGATTATCC AAAACCTCAA
34201 AATGAAGATC TATTAAGTGA ACGCGCTCCC CTCCGGTGGC GTGGTCAAAC TCTACAGCCA
34261 AAGAACAGAT AATGGCATTT GTAAGATGTT GCACAATGGC TTCCAAAAGG CAAACTGCCC
34321 TCACGTCCAA GTGGACGTAA AGGCTAAACC CTTCAGGGTG AATCTCCTCT ATAAACATTC
34381 CAGCACCTTC AACCATGCCC AAATAATTTT CATCTCGCCA CCTTATCAAT ATGTCTCTAA
34441 GCAAATCCCG AATATTAAGT CCGGCCATTG TAAAAATCTG CTCCAGAGCG CCCTCCACCT
34501 TCAGCCTCAA GCAGCGAATC ATGATTGCAA AAATTCAGGT TCCTCACAGA CCTGTATAAG
34561 ATTCAAAAGC GGAACATTAA CAAAAATACC GCGATCCCGT AGGTCCCTTC GCAGGGCCAG
34621 CTGAACATAA TCGTGCAGGT CTGCACGGAC CAGCGCGGCC ACTTCCCCGC CAGGAACCAT
34681 GACAAAAGAA CCCACACTGA TTATGACACG CATACTCGGA GCTATGCTAA CCAGCGTAGC
34741 CCCGATGTAA GCTTGTTGCA TGGGCGGCGA TATAAAATGC AAGGTACTGC TCAAAAAATC
34801 AGGCAAAGCC TCGCGCAAAA AAGCAAGCAC ATCGTAGTCA TGCTCATGCA GATAAAGGCA
34861 GGTAAGTTCC GGAACCACCA CAGAAAAAGA CACCATTTTT CTCTCAAACA TGTCTGCGGG
34921 TTCCTGCATA AACACAAAAT AAAATAACAA AAAAAAAAAA ACATTTAAAC ATTAGAAGCC
34981 TGTNTTACAA CAGGAAAAAC AACCCTTATA AGCATAAGAC GGACTACGGC CATGCCGGCG
35041 TGACCGTAAA AAAACTGGTC ACCGTGATTA AAAAGCACCA CCGACAGTTC CTCGGTCATG

FIG.11A-13

35101 TCCGGAGTCA TAATGTAAGA CTCGGTAAAC ACATCAGGTT GGTTAACATC GGTCAGTGCT
35161 AAAAAGCGAC CGAAATAGCC CGGGGGAATA CATACCCGCA GGCGTAGAGA CAACATTACA
35221 GCCCCCATAG GAGGTATAAC AAAATTAATA GGAGAGAAAA ACACATAAAC ACCTGAAAAA
35281 CCCTCCTGCC TAGGCAAAAT AGCACCCTCC CGCTCCAGAA CAACATACAG CGCTTCCACA
35341 GCGGCAGCCA TAACAGTCAG CCTTACCAGT AAAAAAACCT ATTAAAAAAC ACCACTCGAC
35401 ACGGCACCAG CTCAATCAGT CACAGTGTAA AAAGGGCCAA GTACAGAGCG AGTATATATA
35461 GGACTAAAAA ATGACGTAAC GGTTAAAGTC CACAAAAACC ACCCAGAAAA CCGCACGCGA
35521 ACCTACGCCC AGAAACGAAA GCCAAAAAAC CCACAACTTC CTCAAATCTT CACTTCCGTT
35581 TTCCCACGAT ACGTCACTTC CCATTTTAAA AAAAAACTAC AATTCCCAAT ACATGCAAGT
35641 TACTCCGCCC TAAAACCTAC GTCACCCGCC CCGTTCCCAC GCCCCGCGCC ACGTCACAAA
35701 CTCCACCCCC TCATTATCAT ATTGGCTTCA ATCCAAAATA AGGTATATTA TTGATGATG

FIG.11A-14

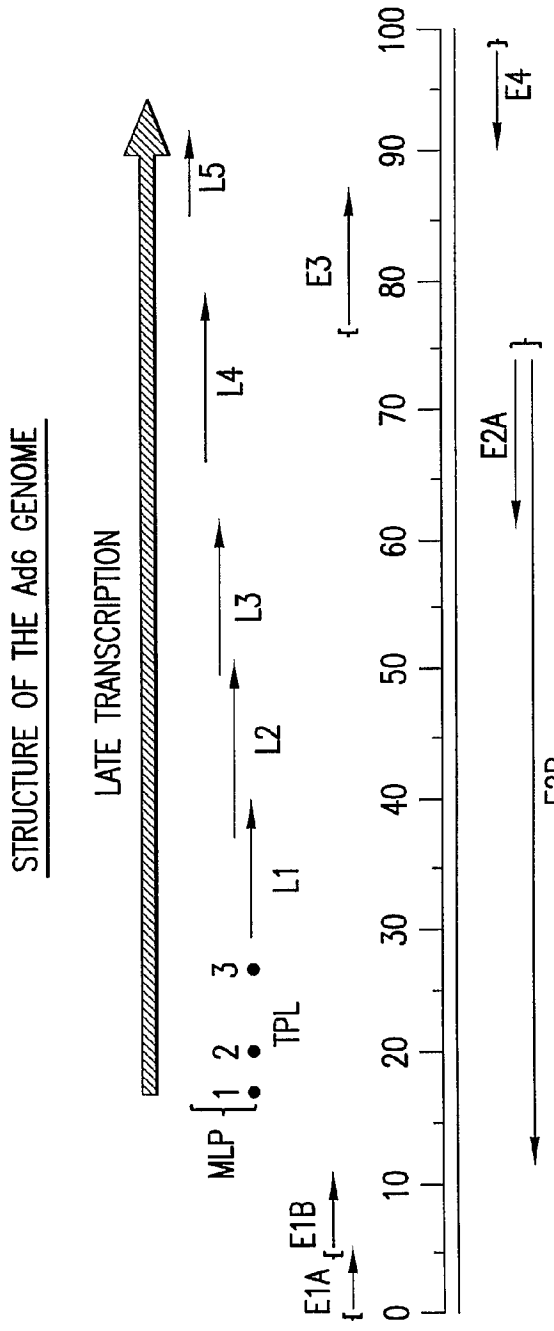

STRUCTURE OF THE Ad6 GENOME
LATE TRANSCRIPTION
MLP
1
2
3
TPL
L1
L2
L3
L4
L5
E1A
E1B
E2A
E2B
E3
E4
0
10
20
30
40
50
60
70
80
90
100

FIG.12

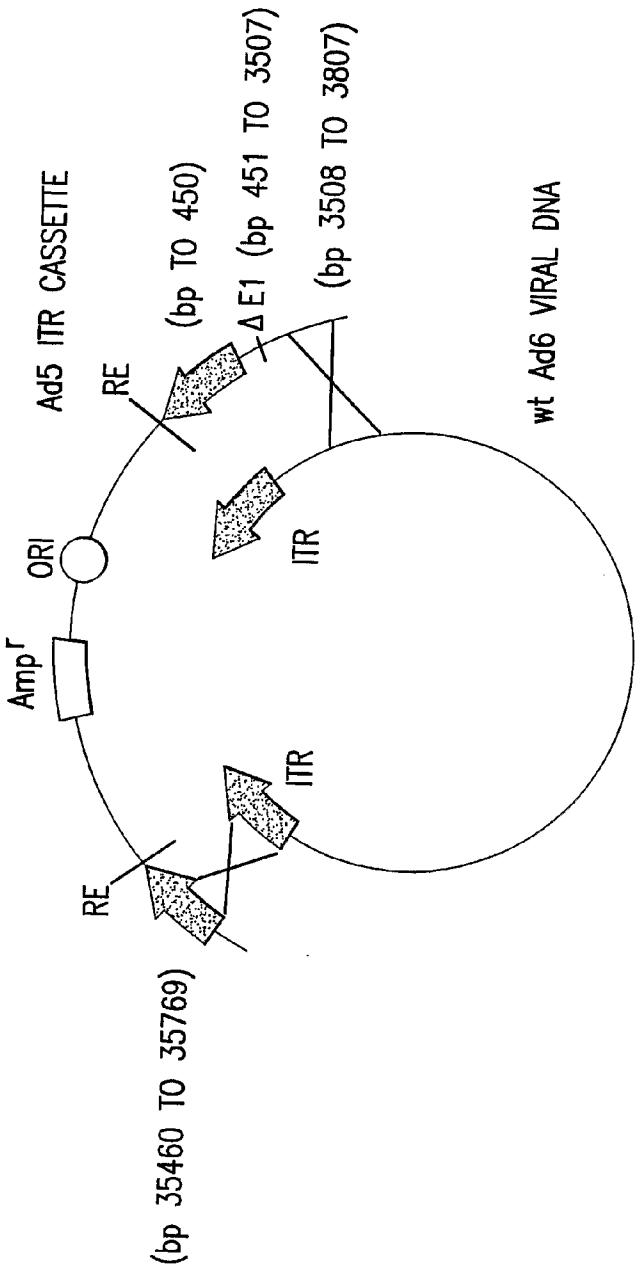

Ad5 ITR CASSETTE
RE
(bp TO 450)
Δ E1 (bp 451 TO 3507)
(bp 3508 TO 3807)
wt Ad6 VIRAL DNA
ORI
Amp^r
ITR
ITR
RE
(bp 35460 TO 35769)

FIG.13